REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-01-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
October 2005	REPRINT	

4. TITLE AND SUBTITLE

Proceedings of the
Quantum Computation for Physical Modeling Workshop 2004,
Old Silver Beach, North Falmouth, MA
12-15 September 2004

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHORS
Editors: Jeffrey Yepez and Timothy Havel*

5d. PROJECT NUMBER
2304

5e. TASK NUMBER
OT

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory /VSBYA
29 Randolph Road
Hanscom AFB, MA 01731-3010

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-VS-HA-TR-2007-1053

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/VSBYA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; distribution unlimited.

13. SUPPLEMENTARY NOTES
Reprinted from *Quantum Information Processing,* Vol. **4**, No. 4, October 2005 © 2005 Springer Science + Business Media, Inc.
*Massachusetts Institute of Technology, Cambridge, MA 02138

14. ABSTRACT

This workshop focused on the development of prototype quantum information processors, quantum algorithms used for physical modeling applications, and efficient numerical simulations.

15. SUBJECT TERMS

Quantum computer, Quantum lattice gas, Quantum simulation, Electron gas, Decoherence, Physical modeling, Physical quantum algorithms, Qubit, Hamiltonian energy, quantum logic gates, Computational complexity

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jeffrey Yepez
UNCL	UNCL	UNCL			19B. TELEPHONE NUMBER *(Include area code)*

Quantum Information Processing, Vol. 4 (2005)

Proceedings of the
Quantum Computation for Physical Modeling Workshop 2004

Editors: Jeffrey Yepez, Timothy Havel

The Air Force Research Laboratory hosted the Quantum Computation for Physical Modeling (QCPM) Workshop 2004 on 12-15 September 2004, our third biannual workshop on quantum computation for physical modeling held at The Harbor View Hotel on Martha's Vineyard. The workshop was hosted by our laboratory and sponsored by the Air Force Office of Scientific Research (AFOSR). The three day workshop comprised six sessions spanning the following research areas: (1) quantum algorithms; (2) quantum effects and devices; (3) superconductive quantum information processing; (4) nuclear magnetic resonance quantum information processing; (5) quantum information, quantum cryptography and quantum communication.

We reviewed the research progress carried out by our extramural university research programs following our laboratorys in-house basic research mission and funded through confederated AFOSR grants. Research articles were presented by government, academic, and commercial researchers, including participants from the following universities: the College of William & Mary, Massachusetts Institute of Technology, Old Dominion University, Texas A&M University, Tufts University, University of California San Diego, University of Florida, and the University of Minnesota. Representatives from four Air Force Research Laboratory directorates, the Air Force Office of Scientific Research, Information, Sensors, and Space Vehicles directorates, were in attendance as well as guest participants from the Army Research Laboratory and the Hewlett Packard Corporation.

Our workshop focused on the development of prototype quantum information processors and quantum algorithms used for physical modeling applications and efficient numerical simulations. The workshop was an outstanding and unqualified success and represents a new milestone in our progress toward achieving a practical quantum computer using state-of-the-art nanotechnology. Additional information about the QCPM Workshop series is located at http://qubit.plh.af.mil.

Quantum Information Processing, Vol. 4 (2005)

Contents to QCPM articles from Volume 4

Quantum Information Processing, Vol. 4, No. 4, October 2005 (© 2005)
DOI: 10.1007/s11128-005-6492-z

Implementation Schemes for the Factorized Quantum Lattice-Gas Algorithm for the One Dimensional Diffusion Equation Using Persistent-Current Qubits

David M. Berns[1,3] and T. P. Orlando[2]

Received January 14, 2005; accepted May 4, 2005

We present two experimental schemes that can be used to implement the Factorized Quantum Lattice-Gas Algorithm for the 1D Diffusion Equation with Persistent-Current (PC) Qubits. One scheme involves biasing the PC Qubit at multiple flux bias points throughout the course of the algorithm. An implementation analogous to that done in Nuclear Magnetic Resonance (NMR) Quantum Computing is also developed. Errors due to a few key approximations utilized are discussed and differences between the PC Qubit and NMR systems are highlighted.

KEY WORDS: Quantum Lattice-Gas; flux qubit; diffusion; quantum computation.

PACS: 03.67.Lx; 85.25.Cp.

1. INTRODUCTION

Most algorithms designed for quantum computers will not best their classical counterparts until they are implemented with thousands of qubits. For example, the factoring of binary numbers with a quantum computer is estimated to be faster than a classical computer only when the length of the number is greater than about 500 digits.[1] Accounting for error correction circuitry[2] would bring the size of the needed quantum computer to be in the thousands of qubits. In contrast, the Factorized Quantum Lattice-Gas Algorithm (FQLGA)[3] for fluid dynamics simulation, even

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
[2]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
[3]To whom correpondence should be addressed. E-mail: dmb@MIT.edu

when run on a quantum computer significantly smaller than the one just discussed, has significant advantages over its classical counterparts.

The FQLGA is the quantum version of classical lattice-gases (CLG).[4] The CLG are an extension of classical cellular automata with the goal of simulating fluid dynamics without reference to specific microscopic interactions. The binary nature of the CLG lattice variables is replaced for the FQLGA by the Hilbert space of a two-level quantum system. The results of this replacement are similar to that of the lattice–Boltzmann model, but with a few significant differences.[5] The first is the exponential decrease in required memory. The second is the ability to simulate arbitrarily small viscosities.

As of today there is a plethora of qubits to choose from when designing a quantum computer, and a promising class is superconducting qubits based on Josepshon junction circuits.[6–10] One major advantage of any of these superconducting systems is the ability to precisely engineer the quantum Hamiltonian, which extends from single qubit design to multi-qubit coupling arrangements to measurement engineering. The quantum computer considered here will be built using the Persistent-Current Qubit (PC Qubit).[6]

The goal of this paper is to show how one can implement a 1D version of the FQLGA with the PC Qubit. To this end we will begin by reviewing the algorithm, specifically the one that simulates the diffusion equation, without a loss of generality in understanding the essence of the algorithm or its general requirements. We will then review the PC qubit and show explicitly how to implement the algorithm with this system. Some important differences between the PC qubit and the two-state system studied in Nuclear Magnetic Resonance Quantum Computation (NMRQC)[11] will be shown to allow for some interesting new techniques in implementing quantum logic. We will also show how to implement the algorithm with the PC qubit in a very analogous way to NMRQC schemes[12] with a few significant differences.

2. FQLGA FOR THE 1D DIFFUSION EQUATION

The first thing one must do in the FQLGA is to define a lattice. Each lattice point n will represent a unique position in the simulated fluid. The simulation will contain a finite number of lattice points, hence space is discretized in the simulation.

Next one must encode the mass density ρ of the fluid at each lattice site. In the FQLGA this is done by building at each lattice site a set $\{i\}$ of coupled qubits. Each qubit represents the motion of particles on the microscopic level in one of a finite set of directions. For the diffusion equation in one dimension, at any point in your fluid, there are only two

possible directions for each particle to be moving, to the left and to the right. Hence, only two qubits are needed to specify the mass density ρ^n at each lattice site. This intuitive reasoning does not extrapolate to higher dimensional simulations because even in two dimensions there would be an infinite number of directions particles could travel in. In higher dimensions one must adhere to much more mathematical conditions to decide on the small set of directions one must include for a faithful simulation.[4] The probability P of a particle to be participating in the motion assigned to each qubit will be encoded in the probability amplitude of the qubit being in its excited state $|1\rangle$. The state of a qubit is thus set to

$$|\Psi_i{}^n\rangle = \sqrt{1 - P_i{}^n}|0\rangle + \sqrt{P_i{}^n}|1\rangle, \tag{1}$$

where i is the qubit index, n is the lattice site index, and $|0\rangle$ is the ground state of the qubit. For the 1D problem considered here, $i = \{1, 2\}$ and $n = \{1, N\}$, where N is the number of lattice sites used in the simulation. One can easily conceive of fluids of multiple phases with multiple types of interactions even in one dimension, in which the size of $\{i\}$ would be much larger, but this will not be considered here. The mass density ρ is then calculated by summing the occupation probabilities for all qubits at a node. At time t=0 in a 1D simulation the occupation probabilities $P_1{}^n$ and $P_2{}^n$ are set to $\rho^n/2$, which is the condition for local equilibrium in the fluid.[13]

Now that the fluid is initialized, one must account for the interaction of particles in the fluid. These collisions are encoded by the application of a unitary transformation to the coupled systems at each lattice site. For the 1D diffusion equation this unitary transformation is

$$\sqrt{\text{swap}} = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1+i & 1-i & 0 \\ 0 & 1-i & 1+i & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}. \tag{2}$$

The basis for computation is the set of four product states: $|0\rangle|0\rangle$, representing no particles at the site, $|0\rangle|1\rangle$, representing the existence of only a particle moving to the right at the site, $|1\rangle|0\rangle$, representing the existence of only a particle moving to the left at the site, and $|1\rangle|1\rangle$, representing particles moving in both directions at the site. To conserve particle number there can be coupling only between the middle two states. The identity transformation on the first and last states corresponds to no collisions and a perfectly elastic collision, respectively. Transformation of the middle two states was something that never existed in the classical algorithm because there was no superposition of these two states.

After collision the states of the qubits at each lattice site are in general entangled, and we denote that state as $|\Upsilon^n\rangle$. The state of each qubit is then measured, and the process described thus far is repeated many times to achieve an ensemble average. Upon completion of these measurements one will have found the post-collision outgoing occupation probabilities, denoted by $P_i{}^n$ once again, and hence the post-collision state of each qubit, denoted by $|\chi_i{}^n\rangle$. Note that the occupation probabilities now represent something very different than before the collision. The particles have now interacted and are ready to move to the next lattice site.

One must now "stream" the occupation probabilities to their new lattice sites. This is done in a classical computer by storing the occupation probabilities at each lattice site that are coming from adjacent lattice sites due to collisions. More precisely, $P_1{}^n$ becomes $P_1{}^{n+1}$ and $P_2{}^n$ becomes $P_2{}^{n-1}$. Periodic boundary conditions are assumed when streaming at the edges of the fluid.

To find the mass density ρ^n at $t=1$ one simply adds the occupation probability for both qubits at site n once streaming has been done. One time step of the algorithm has now been completed. To simulate the next time step simply start the above procedure all over again except now setting the initial states with the new occupation probabilities just found.

The algorithm can be summarized by four major steps, which are illustrated in Fig. 1. The first step encodes the initial state of the fluid by quantum mechanically setting the state $|\Psi_i{}^n\rangle$ of each qubit at each lattice site. The second step transforms the two-qubit product state at each lattice site to in general an entangled state, whose state is denoted by $|\Upsilon^n\rangle$. Third one makes a projective measurement of the post-collision states $|\chi_i{}^n\rangle$, and one must repeat the first three steps to find the outgoing occupation probabilities $P_i{}^n$. In the fourth and final step one streams the mass density with the appropriate post-collision occupation probabilities, from the left with particles representing positive momentum, and from the right with particles representing negative momentum, and the mass density is calculated. Subsequent time steps are identical except for a change in the initial mass density profile, i.e., initial qubit states in the first step.

3. PERSISTENT-CURRENT QUBIT

The fundamental unit of quantum logic we will use to implement the algorithm is the PC Qubit.[14] It consists of a superconducting loop that is interrupted by three Josephson junctions, pictured as x's in Fig. 2(a). The magnetic flux Φ is the only control field for our qubit, and as shown in the figure, is usually denoted by $f=\Phi/\Phi_o$, where $\Phi_o=h/2e$ is a single flux

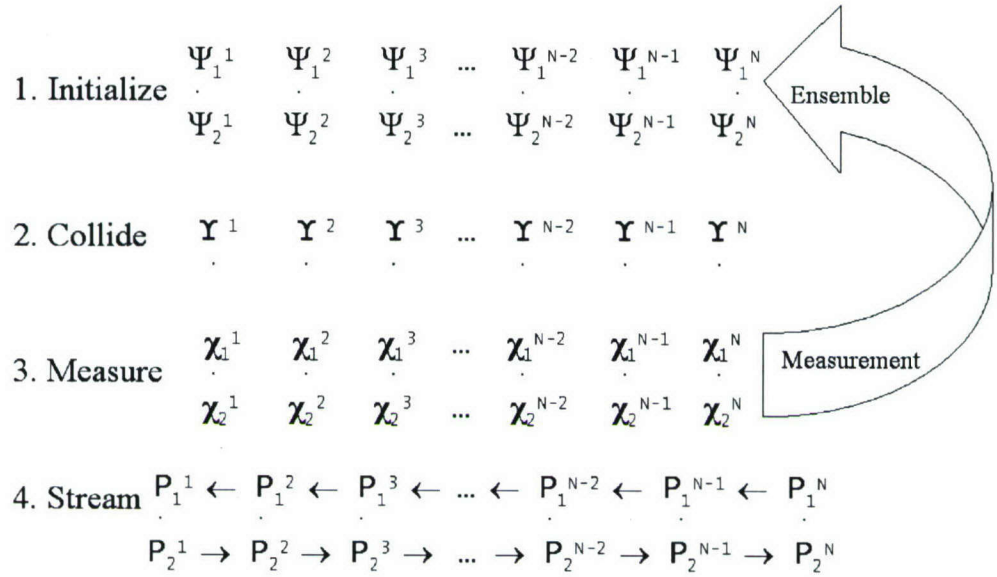


Fig. 1. General summary of the four major steps that comprise one time step of the 1D FQLGA fluid dynamics simulation. The sequence of initialization of mass density, collision of particles, and measurement of post-collision states is repeated many times to make an ensemble measurement. Propagation between collisions is accomplished by storing the adjacent occupation probabilities for a given site in a classical computer, where the mass density is then calculated for this time step. Subsequent time steps utilize these "streamed" occupations when initializing again for the next set of collisions and "streaming".

quantum, h is Planck's constant, and e is the magnitude of the charge of an electron. Physically, a Josephson junction is a small layer of insulator sandwiched between superconductors, so our system is a superconducting loop interrupted by three layers of insulator about 1 nm thick. For single qubit manipulation the magnetic flux through the loop will be modified. The flux seen by a DC SQUID magnetometer, a combination of applied flux and qubit-induced flux, will serve as our measurement variable.

The Hamiltonian of the qubit is derived by considering a circuit element model of our system, which consists of three Josepshon junctions, where two junctions have the same cross-sectional area, and the third is smaller by a factor of α. The constituent relations for an ideal Josephson junction are

$$I = I_c \sin(\varphi), \tag{3a}$$

$$V = \frac{\Phi_o}{2\pi} \frac{\mathrm{d}\varphi}{\mathrm{d}t} \tag{3b}$$

where I is the current through the junction, V the voltage across the junction, I_c the maximum current the junction can hold without a voltage

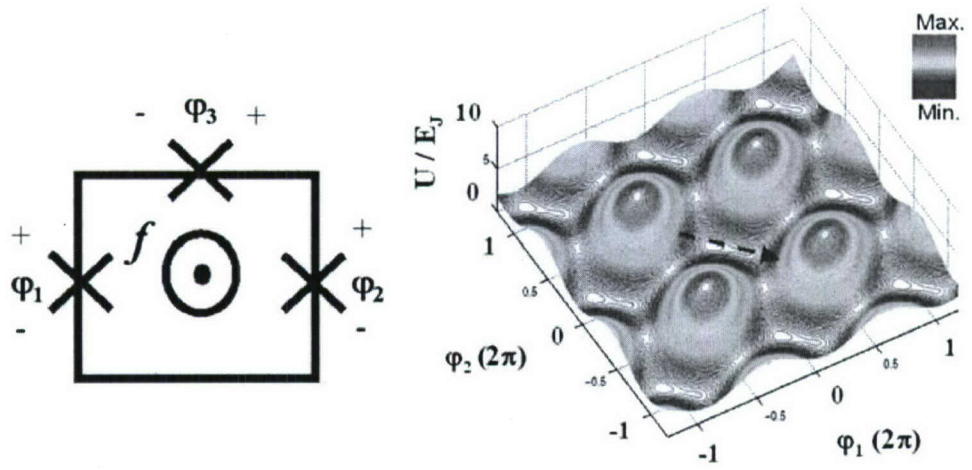


Fig. 2. (a) Schematic drawing of the PC Qubit. The x's represent Josephson junctions, with all connecting leads made of the same superconductor that is part of the junctions. The sign conventions chosen when summing phases are shown, and the magnetic flux penetrating the loop (in units of Φ_o) is labeled by f. (b) The potential energy of the full Hamiltonian for the PC Qubit is plotted when the system is biased at $f=0.495\Phi_o$. The phase particle sees an infinite 2D lattice with unit cells resembling a double well potential.

appearing across it, $\varphi=\theta_1-\theta_2$, and $\theta_{1,2}$ is the gauge-invariant phase that characterizes the superconductor condensate on the $+,-$ side of the junction, respectively. Note that I_c is a function linear in the cross-sectional area of the junction, and hence the third junction has a lower I_c by a factor of α.

The energy associated with an ideal Josephson junction is found by integrating the power from time $t=0$ to some final time t_o, which is equivalent to an integral from zero phase to some phase φ. The energy it takes to set the phase of a Josephson junction to φ is

$$E=\int_o^{t_o}(I_c\sin\varphi')\left(\frac{\Phi_o}{2\pi}\frac{\mathrm{d}\varphi'}{\mathrm{d}t}\right)\mathrm{d}t=\frac{\Phi_o I_c}{2\pi}\int_o^{\varphi}\sin\varphi'\mathrm{d}\phi'=E_j(1-\cos\varphi), \quad (4)$$

where $E_j=\Phi_o I_c/2\pi$.

By including the charging energies due to the capacitance of the junctions, the Hamiltonian of our circuit is[14]

$$H=\frac{P_p^2}{2M_p}+\frac{P_m^2}{2M_m}+E_j[2+\alpha-2\cos(\varphi_p)\cos(\varphi_m)-\alpha\cos(2\pi f+2\varphi_m)], \quad (5)$$

where $\varphi_p=\varphi_1+\varphi_2$, $\varphi_m=\varphi_1-\varphi_2$, $P_p=M_p\mathrm{d}\varphi_p/\mathrm{d}t$, $P_m=M_m\mathrm{d}\varphi_m/\mathrm{d}t$, $M_p=(\Phi_o/2\pi)^2 2C$, and $M_m=(\Phi_o/2\pi)^2 2C(1+2\alpha)$. The number of degrees of

freedom in the problem was reduced by the fluxoid quantization condition[15]

$$\varphi_1 + \varphi_3 - \varphi_2 = 2\pi n + \frac{2\pi\Phi}{\Phi_o}, \tag{6}$$

which forces the sum of the gauge invariant phases to be proportional to the amount of flux quanta modulo an integer multiple of 2π.

We have chosen to associate the capacitive energy, the first two terms in (5), with the kinetic energy, and the ideal Josephson energy, the last four terms in Eq. (5), with the potential energy. The potential energy is that of an infinite lattice of double wells, as seen in Fig. 2(b). The arrow in the plot shows the direction one would take to traverse from one side of a double well to another.

Although quantum mechanics plays a foremost role in deriving the constitutive relations for the superconducting circuit elements, the Hamiltonian for the circuit itself so far has been classical. The quantum version of the circuit can be understood by imagining a phase "particle" in the potential shown in Fig. 2(b). The behavior of this "particle" is analogous to a particle with an anisotropic mass moving in a 2D periodic potential, and so there exist energy bands in a $\vec{k}$-space, which is here related to the charge stored capacitively by the Josephson junctions. By properly choosing E_j/E_c, where $E_c = e^2/2C$, one can remove any $\vec{k}$ (and hence charge) dependence in the energy of the system, and hence can reduce the problem to that of an effective double well. What we have done is choose parameters such that tunneling between adjacent double wells can be neglected relative to the tunneling within a double well in the tight-binding solution(intra-well tunneling typically about 10^4 times more likely), making the solution effectively that of a single double well.

By considering only the lowest two levels of the double well, the equivalent Hamiltonian is

$$\hat{H} = \Phi_o I_p \left(f - \tfrac{1}{2}\right)\hat{\sigma}_z - \tau\hat{\sigma}_x, \tag{7}$$

where $\pm I_p$ are the eigenvalues of circulating current for the two $\hat{\sigma}_z$ eigenstates and τ is the tunneling element from one side of the double well to the other. The energies of the two eigenstates along with a sketch of the double well as a function of applied flux are shown in Fig. 3. One significant difference between this qubit and the one used in NMRQC is the presence of the $\hat{\sigma}_x$ term. The implication of such a term is that the energy of the eigenstates as well as the eigenstates themselves change as the bias field is modified. In Fig. 3 we see that at the classical degeneracy point $f = 1/2$ the qubit's eigenstates are $\hat{\sigma}_x$ eigenstates, while far from $f = 1/2$,

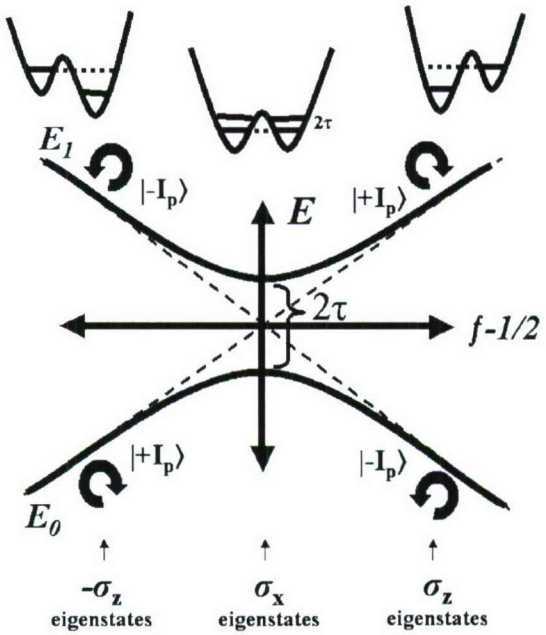


Fig. 3. The energy levels of the PC Qubit are shown as a function of f. The eigenstates of the system change with f and are labeled on the plot. The change in the potential of the phase particle is also depicted at the top of the plot. The energy difference between the states at $f = 1/2$ is seen to be twice the intra-well tunnelling.

but still far from $f = 1$, the eigenstates are those of $\hat{\sigma}_z$. The same thing happens for $f < 1/2$, but now the eigenstates have switched energies, i.e., the ground state here is the first excited state of $\hat{\sigma}_z$ and vice versa.

The PC Qubit has some advantages over other superconducting qubits.[14,16] Charge fluctuations, a consequence of trapped substrate charge, are deemed inconsequential through the choice of parameters used when designing the PC Qubit circuit. Also, flux noise has been reduced in this system over other flux qubits since this system has a smaller loop.

A typical conceptual misconception can be addressed at this point. The two different states used in computation are not related to single Cooper pair behavior. Rather, they are macroscopically distinct states described by the circulating current due to millions of Cooper pairs, characterized by different average induced fluxes when in a magnetic field.

As seen in Sec. 2, the qubits will need to be coupled. For the PC Qubit we have discussed thus far, just as microwaves can only be coupled in through $\hat{\sigma}_z$, coupling between qubits can only be of the form $\hat{\sigma}_z\hat{\sigma}_z$. Slight modifications of the aforementioned qubit design does allow for

different coupling, despite the fact that the structures are still completely planar.[14]

4. IMPLEMENTATION WITH THE PC QUBIT

We now show how one can use the PC Qubit to simulate the 1D diffusion equation. In Sec. 4.1, we elaborate on a scheme based upon changing the flux bias points of the qubits during the algorithm, which will lead to a very general initialization scheme, but a less general collision. In Sec. 4.2, we discuss a more general collision, analogous to that done in NMRQC, and how to initialize the qubits before this general collision. Note that the two implementations differ solely in the way single lattice sites are treated. In both cases one has N qubit pairs fabricated monolithically that are initialized, transformed, and measured simultaneously.

4.1. The Multiple Bias Point Implementation

The first of the four steps of the algorithm is initializing each qubit at each node. As discussed in Sec. 2, each qubit must be initialized into a state of real and positive phase in its own Hilbert space. This set of states consists of all those lying on the real phase geodesic between the ground and first excited states on the Bloch sphere. The ground state of the PC Qubit as a function of applied flux coincidentally also occupies exactly this geodesic on the Bloch sphere, as discussed in Sec. 3. Initialization can thus be accomplished while staying in the ground state by adiabatically changing the applied magnetic flux, as depicted in Fig. 3.

The flux used to set the state of one qubit will be affected by the state of the other qubit and its bias current. This permanent inductive coupling can be accounted for by slightly adjusting the applied flux to compensate for the flux introduced by the other qubit and its bias line. On the other hand, by initializing into the ground state we have avoided the detrimental effects of dephasing and relaxation, as well as the errors found in a typical NMR initialization scheme.[12] We emphasize that the initialization portion of the algorithm is identical for any simulation, whether it be for a different equation, a multi-phase simulation, or in a different number of dimensions.

The second step of the algorithm is the collision. Here we study a very specific unitary transformation, the $\sqrt{\text{swap}}$ described in Sec. 2. This matrix simply "half-way" swaps the middle two (first and second excited) computational states of the coupled system. In NMRQC, the coupled

eigenstates are exactly those computational states, but there are no direct matrix elements connecting these states.[17] When the PC Qubits are coupled, the first and second excited states of the four-level system, denoted as $|1\rangle$ and $|2\rangle$, respectively, are in general not the same as the computational basis states the $\sqrt{\text{swap}}$ intends to affect. However, the DC bias fields of each qubit can be tuned to make these two sets of eigenstates coincide. Once this is done, one can then implement the $\sqrt{\text{swap}}$ by simply oscillating the magnetic field bias at the frequency corresponding to the energy difference between the middle two eigenstates. This is just a Rabi oscillation between the middle two eigenstates, and since one wants to only "half-way" swap the states, the radiation should only be left on for a quarter of a Rabi period.

Besides finding the appropriate bias points such that the middle two eigenstates of the coupled system are very similar to the middle two computational states, one must also verify that the coupling between these states in the presence of an oscillating magnetic field is non-zero. The results of these calculations are shown in Fig. 4. The bias point of qubit 1, f_1, must be chosen to be far from 1/2, but not too far. In these calculations we take $f_1 = 0.508$. In the figure, we see that when qubit 2 is biased at around $f_2 = 0.51$, the first two system excited states are very similar to the middle two computational states, with overlap elements of about 0.97. At this same bias point one sees a Rabi matrix element of about 0.02, which is more than sufficient for our purposes.

This approximate swap has been incorporated into simulation of the FQLGA for the 1D diffusion equation and the results are pictured in Fig. 5. Snapshots of three different times have been shown, for both an ideal simulation and one including the error introduced due to the approximate collision. At time $t = 0$ one can see that we have initialized our fluid to a gaussian profile. Later time steps of the ideal implementation show the expected spreading due to diffusion, while conserving the total number of particles. Increase in the diffusion constant of the approximate collision when compared to the ideal simulation results from the enhanced population in the $|00\rangle$, $|01\rangle$, and $|10\rangle$ states relative to the $|11\rangle$ state due to extra matrix elements in the approximate swap that couple the four states. The matrix elements are actually enhanced more in the upper triangle elements than in the lower triangle elements(with respect to the anti-diagonal), which gives rise to the slight drift to the right that is observed in the simulation.

Even with an ideal swap operator, an interesting timing issue arises upon non-adiabatically switching the bias fields from the initialization settings to the proper settings to do a Rabi oscillation between the two

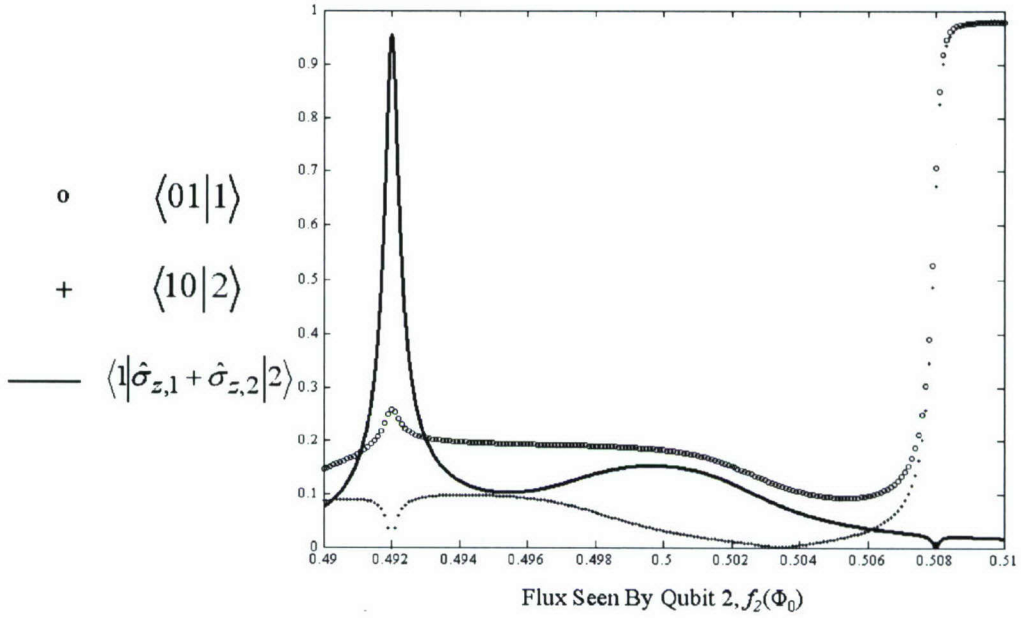


Fig. 4. The overlap between the first (second) excited state $|1\rangle$ ($|2\rangle$) of the PC Qubit coupled system and the $|01\rangle$ ($|10\rangle$) computational state are plotted when qubit 1 is biased at $f = 0.508$. The coupling between $|1\rangle$ and $|2\rangle$ in the presence of an AC magnetic field is also plotted. Qubit coupling equal to τ (same for both qubits) was assumed in the calculation.

middle product states. We first illustrate this timing issue and then show how it can be made negligible by making a larger ensemble measurement.

Once the applied fluxes are changed to those appropriate to perform the approximate swap, the initialized states will most likely not be eigenstates anymore, and hence will begin to precess due to a time-independent perturbation. Assuming things can not be accurately controlled at these timescales, one will have now introduced a random phase difference between the two qubits due to this Larmor precession. This effect is pictured in Fig. 6. The states before the bias fields are switched lie along the same geodesic. Upon changing the magnetic flux seen by each qubit, the qubits begin to precess, out of phase.

The effect of this phase difference δ on the algorithm will be to alter the fraction of particles at each lattice site, post-collision, that are "moving" to the right and to the left. The results of measuring the post-collision occupation probabilities having accounted for a constant phase difference is summarized by

$$P_1 = P_{1,\ \delta=0} + \gamma \sin(\delta), \tag{8a}$$

$$P_2 = P_{2,\ \delta=0} - \gamma \sin(\delta). \tag{8b}$$

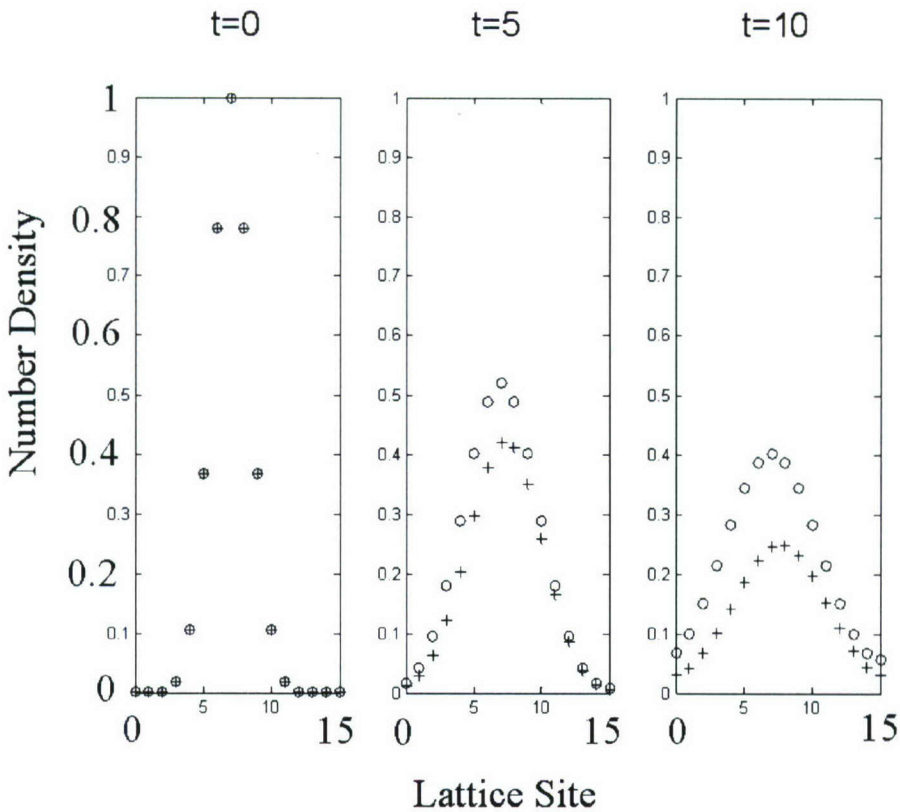


Fig. 5. The results of the FQLGA are simulated having accounted for the approximate nature of the collision proposed in Sec. 4.1 (+). Qubit coupling equal to τ (same for both qubits) was assumed in the approximate swap simulation. The ideal results of the FQLGA are also shown (o).

The effect of this error on the simulation is effectively averaged away when an ensemble is measured, since δ is randomly different for each member of the ensemble. These results are shown in Fig. 7. One can see small random deviations from the ideal simulation that can be made infinitesimally small by measuring a larger ensemble (an ensemble average of 1000 repeated measurements was simulated here).

In summary, an initialization scheme has been developed that is not available to qubits with only one term in their Hamiltonian. This initialization scheme is limited only by the precision of the current source used to create the magnetic field that biases the qubit. The scalability of this scheme relative to those used in NMRQC is an interesting question, but is not resolved here. The collision implementation is also unique to qubits with multiple term Hamiltonians, but the unitary transform implemented is unique to the diffusion equation, and fortuitously simple. A collision

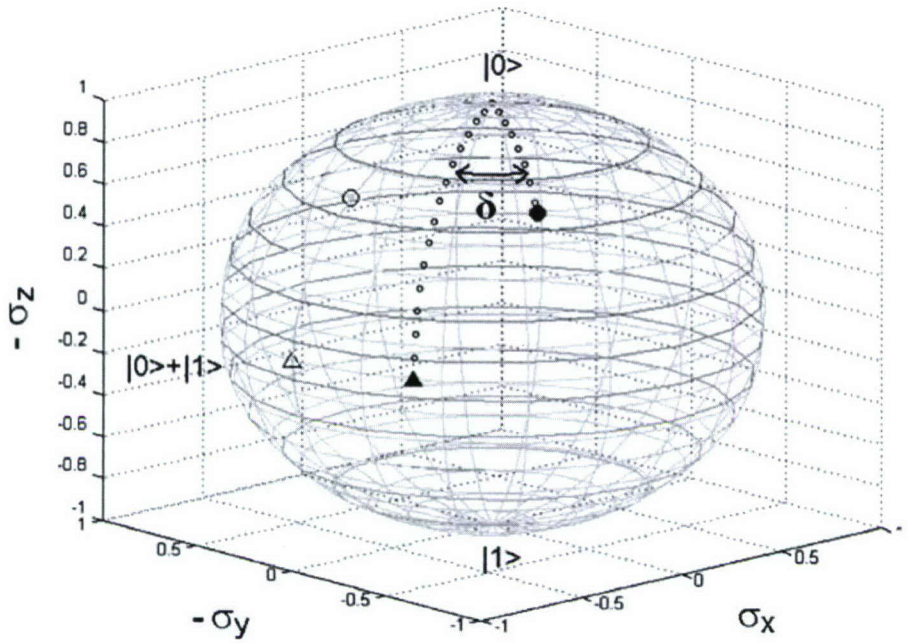


Fig. 6. The unfilled circle and triangle represent two typical initialized states at one lattice point, both on the same north pole to south pole geodesic, before their flux bias is changed to perform the collision. The filled circle and triangle represent the same states after imprecise bias changing has occurred. Imprecisely timed Larmor precession introduces a random phase difference δ between the two states. The unfilled triangle corresponds to the $-\sigma_x$ ground state.

scheme that could be generalized to any unitary transformation would be much more useful.

4.2. Generalized NMRQC-like Implementation

Generalization of the above implementation to any fluid dynamics, i.e., any unitary transformation, can be done in an analogous way to NMRQC schemes. Generalization of the collision transformation consists of using a universal set of quantum computation gates, and decomposing all transformations into a sequence of these.[2] In NMRQC collision is performed by a sequence of single qubit unitary transformations and coupled free evolution. In this section, we will begin by discussing the single qubit rotations needed for a general decomposition, and briefly mention the role they could play in initialization. We will then explore the free evolution of a coupled PC Qubit system, and then show how to combine the single and coupled pulses to implement the collision of the 1D FQLGA for the diffusion equation.

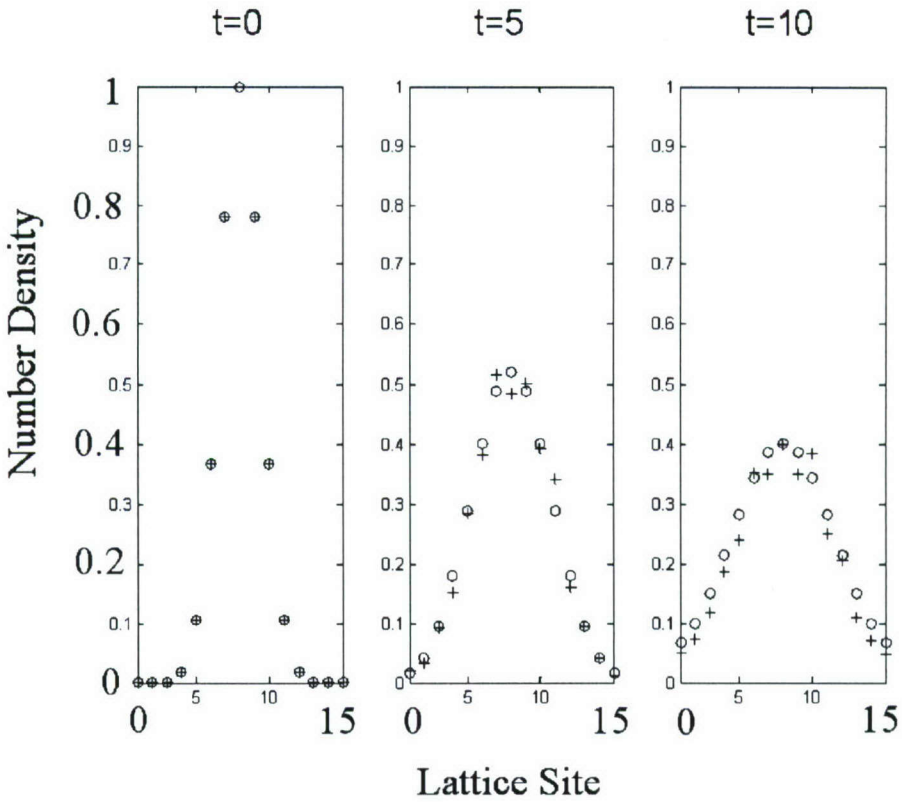


Fig. 7. The results of the FQLGA are simulated having accounted for a random phase difference introduced before the collision for each member of the measurement ensemble (+). The ideal results of the FQLGA are also shown (o).

Single qubit transformations can most easily be achieved in a rotating frame, since here the frequency of precession can be much lower than the Larmor timescale. For this implementation we will only study the case where our qubit is biased at $f = 1/2$. This discussion is easily generalized to any bias point, but the mathematical notation can get quite cumbersome. The Hamiltonian of the PC Qubit in an applied AC field is

$$\hat{H} = w_o \hat{I}_x + g_o \cos(w_o t + \phi)\hat{I}_z, \tag{9}$$

where w_o is the frequency of the applied field, g_o is proportional to the amplitude of the applied field, ϕ is the phase of the applied field, and $\hat{I}_i = \hbar\hat{\sigma}_i/2$.

In the frame rotating about $\hat{I}_x$ this Hamiltonian becomes

$$\tilde{\hat{H}} = \tfrac{1}{2} g_o[\cos(\phi)\hat{I}_z + \sin(\phi)\hat{I}_y]. \tag{10}$$

The quantum state will now precess in this frame about the axis defined by ϕ, with the angle through which the state has precessed given by

$\theta = g_o t/2$. It will be convenient to only consider the set of two rotations defined by $\phi = 0$ and $\pi/2$, which are rotations about $\hat{I}_z$ and $\hat{I}_y$, respectively. These rotations, denoted as $R_z(\theta)$ and $R_y(\theta)$, respectively, can be used in conjunction to bring the qubit state to any point on the Bloch sphere in the rotating frame.

One can use these single qubit rotations not only as part of the collision, but also for initializing, since they can bring the qubit state to anywhere on the Bloch sphere. As already discussed in Sec. 4.1, the ground state of a coupled PC Qubit system is not in general the product of single qubit ground states. Thus, when initializing a qubit via $R_z(\theta)$ and $R_y(\theta)$, one is not starting rotation from the single qubit ground state. However, since the ground state is very close to a product of single qubit ground states, this difference is nearly negligible. In Fig. 8, we show the effects of incorporating this error into the algorithm when the coupling constant is taken to be 1/10 of the qubit resonant frequencies, a rather exaggerated estimate since the coupling is usually much smaller. The diffusion constant is decreased by this approximation, due to the enhanced population in the $|11\rangle$ state relative to the $|00\rangle$ state from the coupling.

The other gate operation needed to form a universal set for general decomposition is coupled free evolution. It is easiest to go to a co-rotating frame where one can have a coupled Hamiltonian only, i.e., no single qubit terms, that is time-independent. In NMRQC this is done by going to the frame where both qubits are rotated around the $\hat{z}$ axis. However, since our coupling does not commute with our single qubit terms, a different method will be used. For notational convenience only, we consider the case where both qubits are biased at $f = 1/2$, where our Hamiltonian is

$$\hat{H} = w_o^1 \hat{I}_x^1 + w_o^2 \hat{I}_x^2 + \frac{2\pi}{\hbar} J_{12} [\hat{I}_z^1 \hat{I}_z^2]. \tag{11}$$

In the co-rotating frame where both qubits are rotating around the $\hat{x}$ axis, one has the Hamiltonian

$$\hat{H} = \frac{\pi}{\hbar} J_{12} [\hat{I}_z^1 \hat{I}_z^2 + \hat{I}_y^1 \hat{I}_y^2] \tag{12}$$

as long as $w_o^1 = w_o^2$. This constraint of $w_o^1 = w_o^2$ imposes limitations on some NMRQC initialization schemes which use frequency selective initialization.

One can now rewrite the unitary collision transformation in the following suggestive way:

$$\sqrt{\text{swap}} = \exp\left[-i\frac{\pi}{8}(\hat{\sigma}_z^1 \hat{\sigma}_z^2 + \hat{\sigma}_y^1 \hat{\sigma}_y^2)\right] \exp\left[-i\frac{\pi}{8}\hat{\sigma}_x^1 \hat{\sigma}_x^2\right]. \tag{13}$$

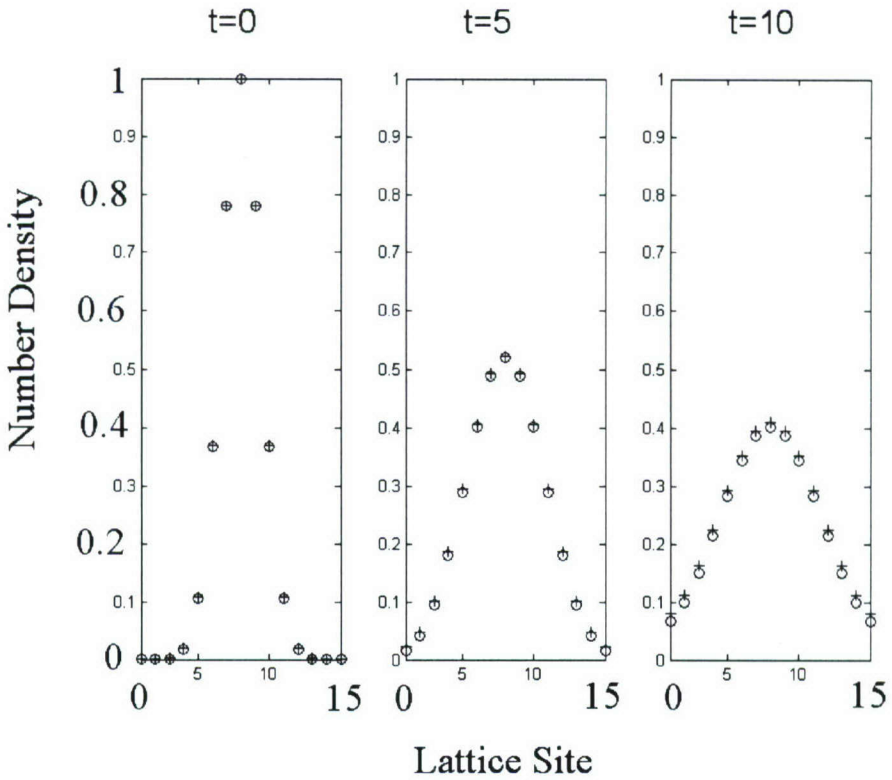


Fig. 8. The results of the FQLGA are simulated for NMR-like single qubit pulse initialization, where errors arise from initializing a coupled ground state that is not a product state (+). The ideal results of the FQLGA are also shown (o).

The first term is just free evolution in the co-rotating frame. The second term can be written as:

$$\exp\left[-i\frac{\pi}{8}\hat{\sigma}_x^1\hat{\sigma}_x^2\right] = R_y^1\left(\frac{\pi}{2}\right)R_y^2\left(\frac{\pi}{2}\right) \times \exp\left[-i\frac{\pi}{8}\hat{\sigma}_z^1\hat{\sigma}_z^2\right]R_y^1\left(-\frac{\pi}{2}\right)R_y^2\left(-\frac{\pi}{2}\right), \tag{14}$$

where the middle term can be written as:

$$\exp\left[-i\frac{\pi}{8}\hat{\sigma}_z^1\hat{\sigma}_z^2\right] = \exp\left[-i\frac{\pi}{8}\left(\hat{\sigma}_z^1\hat{\sigma}_z^2+\hat{\sigma}_y^1\hat{\sigma}_y^2\right)\right]R_z^1(\pi) \times \exp\left[-i\frac{\pi}{8}\left(\hat{\sigma}_z^1\hat{\sigma}_z^2+\hat{\sigma}_y^1\hat{\sigma}_y^2\right)\right]R_z^1(\pi). \tag{15}$$

Hence one can perform a decomposition of the collision transformation into a sequence of single qubit rotations and coupled free evolution.

In summary, we have shown that the PC Qubit can implement the unitary transform that performs collisions in the 1D FQLGA for the

diffusion equation by a single and coupled qubit evolution decomposition. The single qubit rotations were shown to be feasible for qubit initialization as well, with a slight approximation due to the coupled ground state that is not a product state. The coupled free evolution was seen to require identical qubit frequencies over a lattice site, making initialization a bit more challenging.

5. CONCLUSIONS

In this paper, we have shown that the implementation of the FQLGA for the 1D diffusion equation is feasible with PC Qubits. We began by considering the simplest scheme possible using the PC Qubit. This consisted of first initializing the qubits while keeping them in their ground state, and then performing the collision by quickly changing their flux bias points and then performing a single $\pi/2$ pulse. This initialization technique could prove useful, but the way we have implemented the collision is not easily generalized to other collisions. We needed to develop a more general collision scheme, and then see how we could initialize in conjunction with that new scheme.

A more general collision transformation was then discussed by decomposing the unitary matrix into a sequence of single qubit rotations and coupled free evolution. We first developed single qubit rotations for the PC Qubit that could be used as part of the collision decomposition as well as for initializing the occupation probabilities. The initialization was considered only approximate due to the permanent non-commuting coupling between qubits. For the coupled free evolution we saw that transforming to a rotating frame analogously to NMRQC set a strong but feasible constraint on the frequencies of our qubits. Ultimately one would like to remove the constraint of equal frequencies, so that frequency-selective initialization can be done analogously to the NMRQC implementation, alongside the very general collision scheme. One would then also need to account for initialization pulses rotating states from a non-product ground state.

ACKNOWLEDGEMENTS

The authors would like to thank Debra Chen and Jeff Yepez for valuable discussions. This work was supported by the AFOSR/NM grant FA 9550-04-1-0221.

REFERENCES

1. K. Berggren, Quantum Computing with Superconductors, *Proc. IEEE* **92**(10), 1630 (2004).
2. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
3. J. Yepez, *Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, Vol. 1509, (Springer-Verlag, (1999) p. 35.
4. D. A. Wolf-Gladrow, *Lattice-Gas Cellular Automata and Lattice Boltzmann Models*, edited by A. Dold et al. (Springer, Berlin, 2000).
5. J. Yepez, An efficient quantum algorithm for the one-dimensional Burgers equation, quant-ph/0210092.
6. I. Chiorescu et al., Coherent Dynamics of a Superconducting Flux Qubit, *Science* **299**, 1869 (2003).
7. J. R. Friedman et al., Quantum Superposition of Distinct Macroscopic States, *Nature* **406**, 43 (2000).
8. Y. A. Pashkin et al., Quantum Oscillations in Two Coupled Charge Qubits, *Nature* **421**, 823 (2003).
9. J. M. Martinis et al., Rabi Oscillations in a Large Josephon-Junction Qubit, *Phys. Rev. Lett.* **89**, 117901 (2002).
10. D. Vion et al., Manipulating the Quantum State of an Electrical Circuit, *Science* **296**, 886 (2002).
11. I. L. Chuang et al., Bulk Quantum Computation with Nuclear Magnetic Resonance: Theory and Experiment, *Proc. R. Soc. Lond. A* **454**, 447 (1998).
12. M. Pravia et al., Experimental Demonstration of Quantum Lattice Gas Computation, *Quant. Infor. Processing* **2**, 97 (2003).
13. J. Yepez, Quantum Lattice-Gas Model for the Diffusion Equation, *Intl. J. Mod. Phys. C* **12** (9) 1285 (2001).
14. T.P. Orlando et al., Superconducting Persistent-Current Qubit, *Phys. Rev. B* **60**, 15398 (1999).
15. T.P. Orlando and K.A. Delin, *Foundations of Applied Superconductivity* (Addison-Wesley Reading, MA, UK, 1991).
16. D.J. Van Harlingen et al., Decoherence in Josepshon-junction qubits due to critical-current fluctuations, *Phys. Rev. B* **70**, 064517 (2004).
17. G. P. Berman et al., Simulation of the Diffusion Equation on a Type-II Quantum Computer, *Phys. Rev. A* **66**, 012310 (2002).

Quantum Information Processing, Vol. 4, No. 4, October 2005 (© 2005)
DOI: 10.1007/s11128-005-7851-5

On the Interpretation of Energy as the Rate of Quantum Computation

Michael P. Frank[1]

Received January 3, 2005; accepted July 12, 2005

Over the last few decades, developments in the physical limits of computing and quantum computing have increasingly taught us that it can be helpful to think about physics itself in computational terms. For example, work over the last decade has shown that the energy of a quantum system limits the rate at which it can perform significant computational operations, and suggests that we might validly interpret energy as in fact being *the speed at which a physical system is "computing," in some appropriate sense of the word. In this paper, we explore the precise nature of this connection. Elementary results in quantum theory show that the Hamiltonian energy of any quantum system corresponds exactly to the angular velocity of state-vector rotation (defined in a certain natural way) in Hilbert space, and also to the rate at which the state-vector's components (in any basis) sweep out area in the complex plane. The total angle traversed (or area swept out) corresponds to the action of the Hamiltonian operator along the trajectory, and we can also consider it to be a measure of the "amount of computational effort exerted" by the system, or* effort *for short. For any specific quantum or classical computational operation, we can (at least in principle) calculate its* difficulty, *defined as the minimum effort required to perform that operation on a worst-case input state, and this in turn determines the minimum time required for quantum systems to carry out that operation on worst-case input states of a given energy. As examples, we calculate the difficulty of some basic 1-bit and n-bit quantum and classical operations in an simple unconstrained scenario.*

KEY WORDS: Time evolution operator; Margolus–Levitin theorem; Hamiltonian energy; action of the Hamiltonian operator; quantum logic gates; energy as computing; physics as computation; geometric phase; quantum computational complexity.

PACS: 03.65 Ta; 03.67.

[1]Department of Electrical & Computer Engineering, FAMU-FSU College of Engineering, 2525 Pottsdamer St., Rm. 341,Tallahassee, FL 32310, USA. E-mail: mpf@eng.fsu.edu

1. INTRODUCTION

Over the years, the quest to characterize the fundamental physical limits of information processing has also helped to give us a deeper understanding of physics itself. For example, Shannon's studies of the limits of communication[1] taught us that the entropy of a system can also be considered to be a measure of the expected amount of unknown or incompressible information that is encoded in the state of that system. Landauer's[2] and Bennett's[3] analyses of the lower limit to the energy dissipation of computational operations led to Bennett's resolution[4] of the famous Maxwell's demon paradox, via the realization that the demon's record of its past perception is a form of physical entropy, which must be returned to the environment when that information is erased. More recently, Margolus and Levitin[5] showed that the energy of a quantum system limits the rate at which it can perform computational "operations" of a certain type, namely, transitions between distinguishable (orthogonal) quantum states. In the last few years, several articles by Lloyd and colleagues[6–8] have elaborated on this theme by suggesting that we can think of all variety of physical systems (ranging from particles and black holes to the entire universe) as comprising natural computers, with each system's "memory capacity" given by its maximum entropy, and its "computational performance" given by its total energy. We should also note that Ed Fredkin has been promoting a universe-as-computer philosophy for many decades.

The concept of interpreting physics as computing is certainly an exciting theme to pursue, due to its promise of conceptual unification, but we would like to proceed carefully with this program, and take the time to understand the details of this potential unification thoroughly and rigorously. While taking care to get all of the details exactly right, we would like not only to establish that a given physical quantity "limits" or "relates to" a given informational or computational quantity, but also justify the even stronger statement that the physical quantity actually *is*, at root, a fundamentally informational or computational quantity, one that has been traditionally expressed in terms of operationally defined physical units for reasons that can be viewed as being merely historical in nature.

As one the most famous examples of this type of conceptual progression, Rudolph Clausius[9] first defined (differential) entropy as the ratio of differential heat to temperature, $dS = dQ/T$, and at the time, entropy had no further explanation. Later, Ludwig Boltzmann[10] proposed the relation $S \propto -H = \int f \log f \, d\xi$ (where f is a probability density function ranging over particle energies or velocity vectors ξ), which was backed up by his "H-theorem" showing that H spontaneously decreases over time for statistical reasons. In subsequent decades, this relation for entropy evolved

and was generalized to become Boltzmann's eventual epitaph $S = k \log W$, which related entropy to the logarithm of the number of ways W of arranging a system.[11][1] Boltzmann's logarithmic quantity H (in a discrete and negated form) was later recognized by Shannon and others to also be an appropriate measure of the information content of a system. But, Boltzmann's fundamental insight regarding the nature of entropy can be viewed as having gone far beyond just *relating* a physical quantity to an information-based one. Rather, it can be viewed as telling us that physical entropy, at root, *is* really nothing but an informational quantity, one which merely manifests itself in terms of measurable physical units of heat and temperature due to the fact that these quantities themselves have an origin that is ultimately of a statistical nature, e.g., heat as disorganized energy.

Indeed, the long-term quest of physics to eventually create a grand unified "theory of everything" can be viewed as the effort to eventually reveal *all* physical concepts, quantities, and phenomena as being manifestations of underlying structures and processes that are purely mathematical and/or statistical in nature, and that therefore have an informational/computational flavor, at least insofar as the entire realm of formal mathematics can be viewed as being a fundamentally "computational" entity. As one interesting logical conclusion of this conceptual progression, if all observed phenomena are indeed eventually explicable as being aspects of some underlying purely mathematical/computational system, then we can argue that in the end, there really is no need for a separate *physical* ontology at all any more; we could instead validly suppose that the entire "physical" world really *is* nothing but a certain (very elaborate and complex) abstract mathematical or computational object. Such a viewpoint has many attractive philosophical features, at least from the perspective of a hard-core rationalist. One prominent proponent of such musings is Tegmark, e.g., see Ref. 12. Another proposal for unifying mathematics and physics was recently made by Benioff.[13]

However, regardless of one's personal feelings about such far-ranging philosophical agendas, if we can at least show that it is consistent to say that a given physical quantity can be exactly identified with a given mathematical or computational quantity, then, as scientists, we can certainly all agree that the most parsimonious description of physics will indeed be one that does make that identification, since otherwise our description of the world would be burdened with an unnecessary proliferation of artificially distinct concepts, in violation of Ockham's razor, the most fundamental principle of scientific thought.

---

[1]The references to Clausius and Boltzmann in this paragraph are also taken from Ref. 11.

In this paper, we will primarily concern ourselves with just one small aspect of the grander theme of interpreting physics as information processing. Specifically, we focus on the idea of interpreting the physical energy content of a given system as being simply a measure of the rate at which that system is undergoing a certain ubiquitous physical process—namely, quantum state evolution—which can also be viewed as a computational process, as we do in quantum computing. In other words, the premise is that physical energy is nothing but the *rate of quantum computing*, if the meaning of this phrase is appropriately defined. This paper will clarify precisely in what sense this statement is true.

We'll also see that the concept of physical *action*, in a certain (somewhat generalized) sense, corresponds to a computational concept of the *amount of computational effort exerted*, which we'll call *effort* for short.

Of course, it is not necessarily the case that a given system will have been prepared in such a way that all of its physical computational activity will actually be directly applied towards the execution of a target application algorithm of interest. In most systems, only a small fraction of the system's energy will be engaged in carrying out application logic on computational degrees of freedom, while the rest will be devoted to various auxiliary supporting purposes, such as maintaining the stability of the machine's structure, dissipating excess heat to the environment, etc., or it may simply be wasted in some purposeless activity.

For that part of energy that *is* directly engaged in carrying out desired logical operations, we will see that one fruitful application of the computational interpretation of energy will be in allowing us to characterize the *minimum* energy that must be harnessed in order to carry out a given computational operation in a given period of time. In Sec. 12, we will show how to calculate this "difficulty" figure for a variety of simple quantum logic operations, and we briefly discuss how to generalize it to apply to classical reversible and irreversible Boolean operations as well.

2. BACKGROUND

Of course, the earliest hints about the relationship between energy and the rate of computing can be found in Planck's original $E = h\nu$ relation for light, which tells us that an electromagnetic field oscillation having a frequency of ν requires an energy at least $h\nu$, where $h \simeq 6.626 \times 10^{-34}$ J s is Planck's constant. Alternatively, a unit of energy E, when devoted to a single photonic quantum, results in an oscillation (which can be considered to be a very simple kind of computational process) occurring at a cycle rate of $\nu = E/h$.

Also suggestive is the Heisenberg energy–time uncertainty principle $\Delta E \Delta t \geq h/2$, which relates the standard deviation or uncertainty in energy ΔE to the minimum time interval Δt required to measure energy with that precision; the measurement process can be considered a type of computation. However, this relation by itself only suggests that the *spread* or standard deviation of energy has something to do with the rate of a process of interest; whereas we are also interested in finding a computational meaning for the absolute or mean value of the energy, itself.

More recently, in 1992, Tyagi[14] proposed a notion of "computational action" that was based on the amount of energy *dissipated* multiplied by the elapsed time (a quantity which has the same physical units as action) and proposed a theory of optimal algorithm design based on a "principle of least computational action." However, Tyagi's analogy with Hamilton's principle was still a long way from indicating that *physical* action actually *is* computation in some sense, or that physical energy itself (which is, in general, not necessarily dissipated) corresponds to a rate of computation. Still, it was suggestive.

Going much further, in 1998 Toffoli[15] argued that the least-action principle in physics itself can be derived mathematically from *first principles* (rather than as an *ad hoc* physical postulate) as a simple combinatorial consequence of counting the number of possible fine-grained discrete dynamical laws that are consistent with a given macroscopic trajectory. In Toffoli's model, which intriguingly even captures aspects of relativistic behavior, the energy of a state is conjectured to represent the logarithm of the length of its dynamical orbit. Toffoli also gives a correspondence between physical action and amount of computation that is more explicit than Tyagi's, and in which the path with the least Lagrangian action is the one with the greatest amount of "unused" or "wasted" computational capacity. In later papers following up on the present one, we will show that indeed, Lagrangian action corresponds negatively to the portion of the computational effort that does not contribute to an object's active motion.

At around the same time as Toffoli's work, Margolus and Levitin[5] showed that in any quantum system, a state with a quantum-average energy E above the ground state of the system takes at least time $\Delta t \geq t^- = h/4E$ to evolve to an orthogonal state, along with a tighter bound of $\Delta t \geq t_N^- = (N-1)h/2NE$ that is applicable to a trajectory that passes through a cycle of N mutually orthogonal states before returning to the initial state. In the limit as $N \to \infty$, $t_N^- \to h/2E$, twice the minimum time of $t^- = t_2^-$ which applies to a cycle between 2 states. Both bounds are achievable in principle, in freely constructed quantum systems.

In a widely-publicized paper in *Nature* in 2000, Lloyd[6] used the Margolus-Levitin result to calculate the maximum performance of a 1 kg "ultimate laptop," in a hypothetical limiting scenario in which all of

the machine's rest mass–energy is devoted to carrying out a desired computation.

Two years later, Levitin *et al.*[16] investigated the minimum time to perform a specific quantum logic operation, namely a CNOT (controlled-NOT) together with an arbitrary phase rotation, in systems of a given energy E.

In 2003, Giovannetti *et al.*[17,18] explored tighter limits on the time required to reduce the fidelity between initial and final states to a given level, taking into account the magnitudes of both E and ΔE, the system's degree of entanglement, and the number of interaction terms in the system's Hamiltonian.

Results such as the above suggest that energy might fruitfully be *exactly* identified with the rate of raw, low-level quantum-physical "computing" that is taking place within a given physical system, in some appropriate sense, if only the quantity "amount of computing" could be defined accordingly. We would like to show that some well-defined and well-justified measure of the rate at which "computational effort" (not necessarily useful) is being exerted within any quantum system is indeed *exactly* equal to the energy of that system.

3. PREVIEW

In subsequent sections of this paper, we address the aforementioned goal by proposing a well-defined, real-valued measure of the total *amount of change* undergone over the course of *any* continuous trajectory of a normalized state vector along the unit sphere in Hilbert space. This measure is simply given by the line integral of the magnitude of the imaginary component of the inner product between infinitesimally adjacent normalized state vectors along the given path. This quantity is invariant under any time-independent change of basis, since the inner product itself is. As we will show, it is also numerically equal to twice the complex-plane area (relative to the origin) that is circumscribed or "swept out" by the coefficients of the basis vector components, in any basis. For closed paths, this quantity is even invariant under not only rotations but also translations of the complex plane. Finally, our quantity can be perhaps most simply characterized as being *the action of the Hamiltonian* along the path; this is to be contrasted with the usual action (of the Lagrangian), whose precise computational meaning will be addressed in later work.

We propose that the above-described measure of "amount of change" is the most natural measure of the amount of computational *effort* exerted by a physical system as it undergoes a specific trajectory. For any pair of

trajectory endpoints, the effort has a well-defined minimum value over possible trajectories which is obtained along a "geodesic" trajectory between the endpoint states, thereby inducing a natural metric over the Hilbert space.

We will show that in any quantum system, the instantaneous rate at which change occurs (computational effort is exerted) for any state, under any time-dependent Hamiltonian operator, is exactly given by the (Hamiltonian) instantaneous average energy of the state. Thus, the state's energy *is* exactly its rate of computation, in this sense.

We use the word "effort" here rather than "work" both (a) to distinguish our concept from the usual technical meaning of work in physics as being directed energy, and also (b) to connote that effort is something that can be ineffectually wasted; i.e., it does not necessarily correspond to *useful* computational work performed. In fact, we will see that indefinitely large amounts of effort could be expended (inefficiently) in carrying out any given quantum computational task, i.e., in accomplishing a given piece of computational work.

Despite having no upper bound, our concept of effort turns out to still be meaningful and useful for characterizing computational tasks, since (as we will see) any given quantum or classical computational operation does have a well-defined and non-trivial *minimum* required effort for worst-case inputs, which we will call the *difficulty* of the operation. As we will see, for any pair of unitaries U_1, U_2, the difficulty of the operation $U_2 U_1^\dagger$ that takes us from U_1 to U_2 gives a natural distance metric over U_n, the Lie group of rank-n unitary operators.

The difficulty of a computational operation, according to our definitions, determines the minimum time required to perform it on worst-case inputs of given energy, or (equivalently) the minimum worst-case energy that must be devoted to a system in order to perform the operation within a given time. The difficulty thus directly characterizes the computational complexity or "cost" of a given operation, in the same "energy-delay product" units that are popular in electrical engineering, but where the energy here refers to the average instantaneous energy that is *invested* in carrying out the computation, rather than to the amount of energy that is *dissipated.*

4. A SIMPLE EXAMPLE

In this section, we start by presenting a simple, concrete example in order to help motivate our later, more general definitions. Consider any quantum system subject to a constant (time-independent) Hamiltonian operator H. Let $|\mathrm{G}\rangle$ and $|\mathrm{E}\rangle$ be any normalized, non-degenerate pair of the system's energy eigenstates. The labels G and E here are meant

to suggest the ground and excited states of a non-degenerate two-state system, but actually it is not necessary for purposes of this example that there be no additional states of higher, lower, or equal energy.

Since the Hamiltonian is only physically meaningful up to an additive constant, let us adjust the eigenvalue corresponding to vector $|\mathrm{G}\rangle$ to have value 0 (i.e., let $H|\mathrm{G}\rangle = 0$), and then let E denote the eigenvalue of $|\mathrm{E}\rangle$ (i.e., $H|\mathrm{E}\rangle = E|\mathrm{E}\rangle$). For example, for a two-state system, we could let $H = (1+\sigma_z)E/2$ with the usual definition of the Pauli z-axis spin operator $\sigma_z = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$; and let $|\mathrm{G}\rangle = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ and $|\mathrm{E}\rangle = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, thus we have that $H = |\mathrm{E}\rangle\langle\mathrm{E}|$ and so $E = 1$.

Now, consider the initial state $|\psi_0\rangle = (|\mathrm{G}\rangle + |\mathrm{E}\rangle)/\sqrt{2}$ at time $t = 0$, and let it evolve over time under the influence of the system's Hamiltonian, with $|\psi(t)\rangle = \mathrm{e}^{iHt/\hbar}|\psi_0\rangle$ denoting the state vector at time t.[2] Let $c_{|\mathrm{G}\rangle}(t)$ and $c_{|\mathrm{E}\rangle}(t)$ denote $\langle\mathrm{G}|\psi(t)\rangle$ and $\langle\mathrm{E}|\psi(t)\rangle$ respectively, i.e., the components (complex coefficients) of the state vector $|\psi(t)\rangle$ when decomposed in an orthonormal basis that includes $|\mathrm{G}\rangle$, $|\mathrm{E}\rangle$ as basis vectors.

Initially, $c_{|\mathrm{G}\rangle}(t) = c_{|\mathrm{E}\rangle}(t) = 1/\sqrt{2}$. Over time, $c_{|\mathrm{E}\rangle}$ phase-rotates in the complex plane in a circle about the origin, at an angular velocity of $\omega_{|\mathrm{E}\rangle} = E/\hbar$. In time $t = 2E/h$, it rotates by a total angle of $\theta = \pi$. The area swept out by the line between $c_{|\mathrm{E}\rangle}(t)$ and the origin is $a_{|\mathrm{E}\rangle} = \frac{1}{2}\pi|c_{|\mathrm{E}\rangle}|^2 = \pi/4$. This is the area of a semi-circular half-disk with radius $r_{|\mathrm{E}\rangle} = |c_{|\mathrm{E}\rangle}| = 1/\sqrt{2}$. Meanwhile, $c_{|\mathrm{G}\rangle}(t)$ is stationary and sweeps out zero area. The total area swept out by both components is thus $a = \pi/4$. This evolution is depicted in Fig. 1.

Does the area swept out by the complex components of the state vector depend on the choice of basis? We will answer this question in a much more general setting later, but for now, consider, for example, a new basis that includes basis vectors $|0\rangle$, $|1\rangle$ where $|0\rangle = (|\mathrm{G}\rangle + |\mathrm{E}\rangle)/\sqrt{2}$ and $|1\rangle = (|\mathrm{G}\rangle - |\mathrm{E}\rangle)/\sqrt{2}$. Consider the evolution again starting from the same initial state as before, $|\psi_0\rangle = |0\rangle$. Note that the final state after time $t = 2E/h$ is $|1\rangle$. In the new basis, the coefficients $c_{|0\rangle}(t)$ and $c_{|1\rangle}(t)$ respectively trace out the upper and lower halves of a circle of radius $1/2$ centered at the point $1/2 + \mathrm{i}0$. The total area swept out by both components (on lines between them and the origin) is the area of this circle, namely $a = \pi(1/2)^2 = \pi/4$. (See Fig. 2.) Note that the total area in this new basis is still $\pi/4$.

At this point we may naturally ask, is the area the same in *any* fixed basis? Later we will show that the answer is yes; in general, the area swept

[2]For convenience, we use the opposite of the ordinary sign convention in the time-evolution operator.

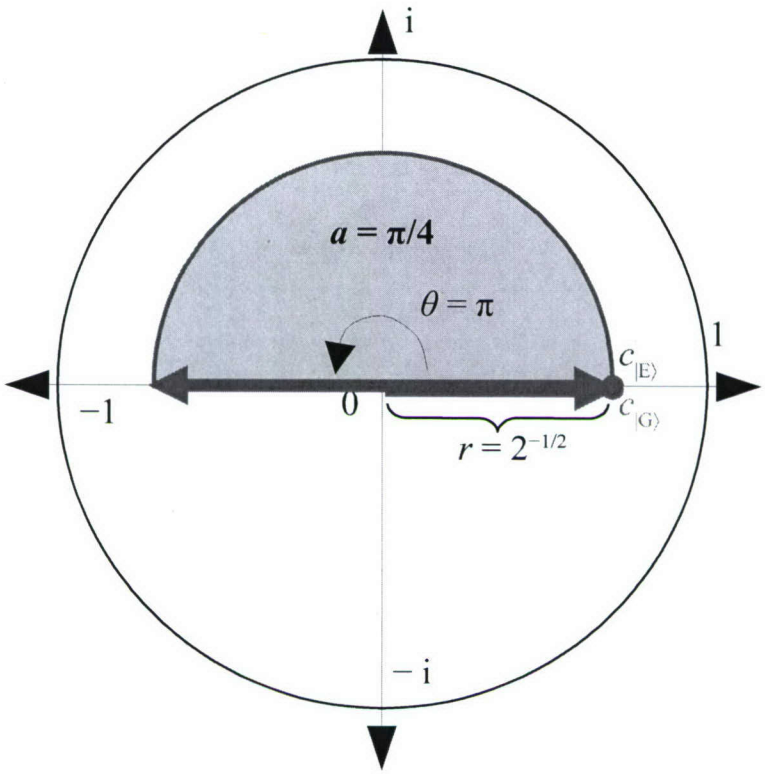


Fig. 1. Under the Hamiltonian $H = E|\mathrm{E}\rangle\langle\mathrm{E}|$, starting from the initial state $|\psi_0\rangle = (|\mathrm{G}\rangle + |\mathrm{E}\rangle) \cdot 2^{-1/2}$, the complex coefficient $c_{|\mathrm{E}\rangle} = \langle\mathrm{E}|\psi(t)\rangle$ of $|\mathrm{E}\rangle$ (the excited state) in the superposition sweeps out a half-circle in the complex plane with area $\pi/4$ in time $t = 2E/h$, while the ground-state coefficient $c_{|\mathrm{G}\rangle}$ remains stationary.

out is independent of the basis for *any* trajectory of *any* initial state. The area swept out will be (proportional to) our proposed measure of the amount of computational effort exerted by a system in undergoing any specific state-vector trajectory.

5. GENERAL FRAMEWORK

In this section we proceed to set forth the general mathematical definitions and notations to be used in the subsequent analysis.

5.1. Time-independent Case

Let $\mathcal{H}$ be any Hilbert space. Any linear, norm-conserving, invertible, continuous and time-independent dynamics on such a space must proceed via the application of a unitary time-evolution operator, expressible as

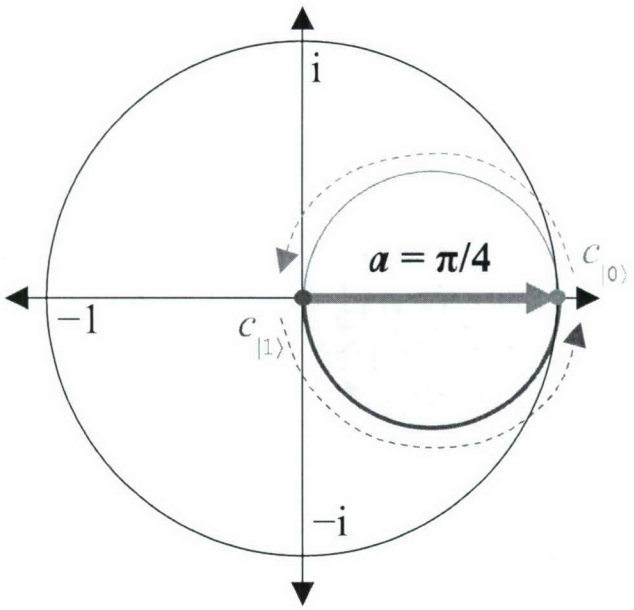


Fig. 2. The evolution from Fig. 1, re-plotted in the basis $|0\rangle = (|G\rangle + |E\rangle) \cdot 2^{-1/2}$, $|1\rangle = (|G\rangle + |E\rangle) \cdot 2^{-1/2}$. The coefficients of $|0\rangle$ and $|1\rangle$ together sweep out a full circle, but the total area swept out is still $\pi/4$.

$$U = U(\Delta t) = e^{iA(\Delta t)} = e^{iH\Delta t}, \tag{1}$$

where Δt is the length of a given time interval, $A(\Delta t) = H\Delta t$ maps the interval to an Hermitian operator A that is proportional to Δt, and H is an Hermitian operator with units of angular frequency. For any two times $t_1, t_2 \in \mathbb{R}$, and for any initial state vector $|\psi\rangle = |\psi(t_1)\rangle$ at time t_1, the implied state at any other time t_2 is given by $|\psi(t_2)\rangle = U(\Delta t)|\psi(t_1)\rangle$, where $\Delta t = t_2 - t_1$. We will sometimes also write U and A as functions of the directed pair of times, written $t_1 \to t_2$. We will sometimes call the U and A operators "cumulative" when the interval Δt is not infinitesimal.

Note that in Eq. (1) we are using the opposite of the usual (but arbitrary) negative-sign convention in the exponent; this is an inessential but convenient choice, in that later it will let us automatically associate positive energies with positive (i.e., counter-clockwise) phase velocities for the coefficients of state components.

For convenience, for any operator O and vector v, we will sometimes use the notation $O[v]$ as an abbreviation for the expectation value $\langle v|O|v\rangle$.

Now, of course, the eigenvectors of U are also eigenvectors of A and H, so H's expectation value $H[\psi]$ for any initial vector $\psi(t_1) \in \mathcal{H}$ is

preserved by the time-evolution $\psi(t_1) \to \psi(t_2)$. This conserved quantity (whose existence follows from time-independence even more generally, via Nöther's theorem) is called the *Hamiltonian energy* of the system. Although in our expressions it has the dimensions of angular velocity, this is the same as energy if we choose units where $\hbar = 1$, as is customary. Thus, H is called the Hamiltonian operator. We will call the operator $A = A(t_1 \to t_2)$ the *cumulative action of the Hamiltonian from time* t_1 *to* t_2, where some of the qualifying phrases may be omitted for brevity. The reasons for the use of the word "action" will be discussed later.

For convenience in the subsequent discussions, we will often just set $t_1 = 0$ (without loss of generality) and write $U = U(t) = U(0 \to t) = e^{iHt}$. We refer to the complete operator-valued function $\lambda t.U(t)$ for all t values in some range (which usually includes $t = 0$, for which $U(0) = I$) as a *unitary trajectory* over that time interval. Also, for any t we write $A(t) :\equiv A(0 \to t)$ for the cumulative action from 0 to t.

Differentiating $U(t)$ with respect to time and applying the result to an initial state $|\psi(0)\rangle$ then yields us Schrödinger's equation in various forms that we'll use,

$$\dot{U} = \frac{dU(t)}{dt} = \frac{d}{dt}e^{iHt} = iHe^{iHt} = iHU(t) \tag{2}$$

$$\frac{d}{dt}U(t)|\psi(0)\rangle = iHU(t)|\psi(0)\rangle \tag{3}$$

$$|\dot{\psi}\rangle = \frac{d}{dt}|\psi(t)\rangle = iH|\psi(t)\rangle \tag{4}$$

$$\frac{d}{dt} = iH, \tag{5}$$

where again, note that we are using $\hbar = 1$ and the opposite of the usual sign convention. Note also that we are able to differentiate e^{iHt} in Eq. (2) because d/dt commutes with H, since H here is a constant.

5.2. Time-dependent Case

The natural generalization of Eq. (5) (the operator form of Schrödinger's equation) to a system with a time-dependent Hamiltonian $H(t)$ is of course just

$$\frac{d}{dt} = iH(t), \tag{6}$$

where now $H(t)$ is permitted to vary over time, though often with a constraint that it be differentiable, smooth, or analytic.

One may at first think that in this time-dependent context, we could appropriately generalize the time-evolution operator equation (1) by simply changing the definition of the action operator A (as a function of t) from the original $A(t) = Ht$ to what one might naïvely think would be the obvious generalization to a time-dependent H,

$$A(t) = \int_{\tau=0}^{t} H(\tau) \mathrm{d}\tau, \tag{7}$$

while still keeping the relation $U(t) = \mathrm{e}^{\mathrm{i}A(t)}$. But in fact, the definition (7) does not work for this purpose, since in general the values of $H(\tau)$ at different times τ will not commute with each other; taking the integral loses all information about their relative time-ordering, and the time-derivative of $U(t)$ will no longer be equal to $\mathrm{i}H(t)$ as required, since $\mathrm{d}/\mathrm{d}t$ will no longer commute with $H(t)$.

The standard way to repair this problem (discussed in almost any quantum field theory textbook, e.g., Ref. 19) is to define a time-ordering meta-operator $\mathcal{T}$, which takes a given operator expression and reorders its internal operator products so that operators associated with earlier time points are applied first in all products (reading right-to-left). For example, as a matter of definition,

$$\mathcal{T}[H(t_1)H(t_2)] :\equiv \begin{cases} H(t_1)H(t_2) & \text{if } t_1 > t_2 \\ H(t_2)H(t_1) & \text{otherwise} \end{cases} \tag{8}$$

With this notational convention, we can write

$$U(t) = \mathcal{T}\mathrm{e}^{\mathrm{i}A(t)}, \tag{9}$$

where $A(t)$ is as defined in Eq. (7), and the meaning of this meta-expression will be well-defined and consistent with Eq. (6) applied to $U(t)$. But the problem with this approach is that the expression $A(t)$ in (9) no longer denotes a "first class object" of our language, but rather is a sort of meta-mathematical place-holder to be manipulated via a rather complex interpretational procedure, which involves applying Eq. (8) to uncountably many infinitesimal pieces of the integrals appearing in the Taylor-expanded version of Eq. (9). There is no longer any simple, direct relationship between the properties of the linear operator $A(t)$ defined in Eq. (7) (e.g., its eigenvalues and eigenvectors) and the properties of $U(t)$.

Thus, in what follows we will find it more useful to instead abandon Eq. (7), and take the rather more concrete approach of simply redefining $A(t)$ for a given unitary trajectory $U(t)$ to be the unique continuously time-dependent Hermitian operator such that $A(0) = 0$ and

$$U(t) = \mathrm{e}^{\mathrm{i}A(t)} \tag{10}$$

(with *no* time-ordering operator!) for all t. To see that such an A indeed exists and is unique, note that since each particular $U = U(t)$ (at a given moment) is unitary, it is a normal operator and can thus be given a spectral decomposition

$$U = \sum_i u_i |u_i\rangle\langle u_i| \tag{11}$$

where $\{|u_i\rangle\}$ and $\{u_i\}$ respectively comprise an orthonormal eigenbasis of U and the corresponding unit-modulus eigenvalues. We can therefore define the multi-valued logarithm of U by

$$\begin{aligned} \ln U &= \ln \sum_i u_i |u_i\rangle\langle u_i| \\ &:\equiv \sum_i (\ln u_i)|u_i\rangle\langle u_i| \\ &= \sum_i \mathrm{i}\arg(u_i)|u_i\rangle\langle u_i| \qquad (12) \\ &= \sum_i \mathrm{i}[\mathrm{Arg}(u_i) + 2\pi n_i]|u_i\rangle\langle u_i|, \qquad (13) \end{aligned}$$

where in step (12) we have used the fact that $|u_i| = 1$, and where in line (13) $\mathrm{Arg}(u_i) \in [0, 2\pi)$ denotes the principal value of the multivalued function $\arg(u_i)$, while the n_i values may be any integers. Although we see that there are infinitely many values of $(\ln U)$ for any individual U in isolation, nevertheless there *is* a unique single-valued definition of the entire function $L(t) = \ln U(t)$, given the function $U(t)$, that is *continuous* over t and where $L(0) = 0$.

The uniqueness is due to the fact that $U(t)$ varies continuously in t, and thus, if we like, the eigenbasis $\{|u_i(t)\rangle\}$ that we choose for U at each moment (which has k free gauge-like parameters determining the u_i, where $k = \dim \mathcal{H}$) can vary continuously as well. Given basis vectors $|u_i\rangle$ (and thus u_i values) that change continuously, it follows that at any moment, only one assignment of values to the n_i parameters can possibly yield continuity with the logarithm value $L(t - \mathrm{d}t)$ at the previous moment, since any other choice would (discontinuously) change one of the phase angles $\mathrm{Arg}(u_i) + 2\pi n_i$ in the expression (13) by an amount that is (infinitesimally close to) a multiple of 2π. The n_i parameters can (and must) change by ± 1 from their preceding values (while leaving $L(t)$ continuous) only at a discrete set of time points, namely those where the continuously changing u_i value crosses the branch cut of the Arg() function (in some direction), and $\mathrm{Arg}(u_i)$ jumps by $\mp 2\pi$.

Now, given this uniquely defined unitary trajectory logarithm $L(t) = \ln U(t)$, we simply define our action operator as $A(t) = -\mathrm{i}L(t)$, and then trivially we have that $U(t) = \mathrm{e}^{\mathrm{i}A(t)}$ holds for all t, where the exponential can be defined via the spectral decomposition of A (equivalently to the standard Taylor-series definition), thereby inverting the logarithm.

Meanwhile, the entire unitary trajectory $U(t)$ itself is derived from the Hamiltonian trajectory $H(t)$ by setting $U(0) = I$ and applying the operator form (6) of the time-dependent Schrödinger equation to $U(t)$. So $(\mathrm{d}/\mathrm{d}t)U(t) = \mathrm{i}H(t)U(t)$, and we are thereby guaranteed that in fact

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\mathrm{i}A(t)} = \mathrm{i}H(t)\mathrm{e}^{\mathrm{i}A(t)} \tag{14}$$

as desired, which (recall) failed to be true (in the absence of a time-ordering operator) for the $A(t)$ defined in Eq. (7).

For reasons we will explain, we will refer to a complete function $\lambda t\,.\,A(t)$ as defined by Eq. (10) as the *cumulative Hamiltonian action trajectory* implied by the Hamiltonian trajectory $H(t)$.

In cases where $H(t) = H$ is constant over time, note that this definition of $A(t)$ reduces to the simple Ht form that we used back in Eq. (1). This follows from the observation that the definition $A(t) = Ht$ indeed solves Eq. (10) when H is constant, and the fact that (as we just showed) the $A(t)$ implied by Eq. (10) is unique under the continuity constraint.

Later, we will see the importance of the Hamiltonian action trajectory $A(t)$, and discuss the precise meaning and computational interpretation of its expectation value when applied to a given state.

To clarify our terminology, note that in this document we are using the word *action* in a somewhat more general sense than is usual; typically in physics (e.g., in Hamilton's principle) "action" just refers to the quantity having units of action that is obtained by integrating the Lagrangian $L = pv - H$ along some path. However, it is also perfectly valid and reasonable to consider the more general notion of the action that is associated with *any* quantity that has units of energy, by setting the time-derivative of that action along some path to be equal to that energy.

Indeed, we will see later that the time-derivative of the cumulative Hamiltonian action $A(t)$ (as we have defined it) along a given trajectory is in fact exactly the instantaneous Hamiltonian energy $H(t)$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}A(t)[\psi(0)] = H(t)[\psi(t)], \tag{15}$$

similarly to how the time-derivative of the ordinary (i.e., Lagrangian) action along a given trajectory is the instantaneous Lagrangian energy $L(t)$.

As a final piece of notation which will help us generalize our results to the time-dependent case, we will sometimes write $U'(t)$ to refer to the "instantaneous" unitary transformation that applies over an infinitesimal time interval $\mathrm{d}t$ at time t, that is,

$$\begin{aligned} U'(t) &:\equiv U(t \to t+\mathrm{d}t) \\ &= 1+\mathrm{i}H(t)\mathrm{d}t. \end{aligned} \tag{16}$$

Note also that any larger transformation $U(t_1 \to t_2)$ can be expressed as the time-ordered product of all the infinitesimal $U'(t)$ over the continuum of times t in the range from t_1 to t_2. That is, we can write

$$U(t_1 \to t_2) = \mathcal{T} \prod_{t=t_1}^{t_2} U'(t) \tag{17}$$

with the opposite ordering if $t_2 < t_1$. Thus, $U'(t)$ uniquely defines $U(t)$, so we will sometimes refer to $U'(t)$ as the unitary trajectory also.

We should keep in mind that although the complete unitary trajectory $U(t)$ (or $U'(t)$) between t_1 and t_2 determines the overall transformation $U(t_1 \to t_2)$, the converse is not true: Knowing the cumulative $U = U(t_1 \to t_2)$ for a particular pair of times t_1, t_2 is of course insufficient to determine a unique unitary trajectory $U(t)$, since in general infinitely many cumulative action operators $A = A(t_1 \to t_2)$ can exponentiate to yield the same cumulative U (since expression (13) is multivalued), and furthermore, in the time-dependent case, a continuum of different Hamiltonian trajectories $H(t)$ (which determine $U'(t)$) could implement a given cumulative action operator A.

We will similarly use the notation $A'(t) = H(t)\mathrm{d}t$ to denote the infinitesimal action operator that applies from time t to $t+\mathrm{d}t$; note that $U'(t) = \mathrm{e}^{\mathrm{i}A'(t)} = 1+\mathrm{i}H(t)\mathrm{d}t$.

6. DEFINING COMPUTATIONAL EFFORT

With the above general definitions and observations aside, let us now proceed to define our concept of the amount of computational effort exerted by a system in undergoing a state trajectory $|\psi(t)\rangle$ between two times.

We will find it easiest to define this quantity first for the case of a system with a time-independent Hamiltonian $H(t) = H = \text{const}$. Later, we will show how our results can be generalized to the time-dependent case.

Let $|v\rangle$ be any eigenvector of H, and ω the corresponding eigenvalue, which is real since H is Hermitian. That is, let $H|v\rangle = \omega|v\rangle$. Thus, $|v\rangle$ is also an eigenvector of the cumulative action operator $A(t) = Ht$ for any t, with eigenvalue $\alpha = \omega t$.

First, when t is an infinitesimal $\mathrm{d}t$, consider the instantaneous $U' = 1 + iH\mathrm{d}t$. Clearly, $|v\rangle$ is an eigenvector of U', since $U'|v\rangle = (1 + iH\mathrm{d}t)|v\rangle = (1 + i\omega\mathrm{d}t)|v\rangle = u|v\rangle$, where the scalar $u = 1 + i\omega\mathrm{d}t = e^{i\omega\mathrm{d}t} = e^{i\mathrm{d}\alpha}$. Thus, under application of U', the eigenvector $|v\rangle$ transforms to $|v'\rangle :\equiv e^{i\omega\mathrm{d}t}|v\rangle = e^{i\mathrm{d}\alpha}|v\rangle$, that is, it phase-rotates in the complex plane at angular velocity ω through an infinitesimal angle $\mathrm{d}\alpha$. Note also that

$$\begin{aligned} \Im\langle v|v'\rangle &= \Im\langle v|(1 + i\mathrm{d}\alpha)|v\rangle = \Im(1 + i\mathrm{d}\alpha)\langle v|v\rangle \\ &= \mathrm{d}\alpha = \langle v|\omega\mathrm{d}t|v\rangle = \langle v|A'|v\rangle = A'[v]. \end{aligned} \tag{18}$$

That is, when $|v\rangle$ is an eigenvector of H, the magnitude of the imaginary part of the inner product between infinitesimally adjacent state vectors is equal to the expectation value $A'[v]$ of the infinitesimal action operator $A' = H\mathrm{d}t$ applied to the state. As we go on, we will extend the relationship (18) to non-infinitesimal trajectories, non-eigenvectors, and time-dependent Hamiltonians.

Next, note that the eigenvectors $|v\rangle$ of H are also eigenvectors of the cumulative action operators $A(t) = Ht$ and cumulative unitaries $U(t) = e^{iA(t)} = e^{iHt}$, and vice-versa. Let $A(t)|v\rangle = \alpha(t)|v\rangle$, with $|v\rangle$ a fixed eigenket of $A(t)$, and with $\alpha(t) = \omega t$ as its eigenvalue. Then, $U(t)|v\rangle = e^{iA(t)}|v\rangle = e^{i\alpha(t)}|v\rangle = u(t)|v\rangle$ where $u(t) = e^{i\alpha(t)}$. Thus, upon the application of U, $|v\rangle$ gets multiplied by the phase factor $u(t)$, or (we can say) rotated by a total phase angle of $\alpha(t) = \omega t$, which could be much greater than 2π in long evolutions, as can also be seen by integrating $\mathrm{d}\alpha$ over t. Note also that if we integrate $\Im\langle v|v'\rangle$ along the trajectory, we still get the cumulative action $A(t)[v(0)]$:

$$\int_{\tau=0}^{t} \Im\langle v(\tau)|v'(\tau)\rangle = \int_{\tau=0}^{t} \Im\langle v(\tau)|(1 + i\omega\mathrm{d}\tau)|v(\tau)\rangle \tag{19}$$

$$= \omega t = \alpha(t) = \langle v(0)|A(t)|v(0)\rangle. \tag{20}$$

Next, consider an arbitrary pure state $|\psi(0)\rangle = \sum_i c_i(0)|v_i\rangle$, where the $|v_i\rangle$ are normalized eigenstates of H with eigenvalues ω_i, and the $c_i(0)$ are the initial coefficients of the $|v_i\rangle$ in the superposition. The state at time t can be expressed as

$$|\psi(t)\rangle = \sum_i \exp[i\alpha_i] c_i(0)|v_i\rangle$$
$$= \sum_i \exp[i\omega_i t] c_i(0)|v_i\rangle$$
$$= \sum_i c_i(t)|v_i\rangle, \tag{21}$$

where we see that each coefficient $c_i(t) = \exp[i\omega_i t]c_0(t)$ (in the fixed basis $\{|v_i\rangle\}$) simply phase-rotates with angular velocity ω_i along an origin-centered circle in the complex plane with constant radius $r_i = |c_i|$. Over any amount of time t, we see that c_i rotates in the complex plane by a total angle of $\alpha_i = \omega_i t$, while the line in the complex plane that joins c_i to the origin sweeps out an arc with an area of $a_i = \frac{1}{2}\omega_i t r_i^2$. (See Fig. 3 for an illustration of the area swept out in the infinitesimal case.) For example, in time $t = 2\pi/\omega_i$, coefficient c_i sweeps out a complete disk of area $a_i = \pi r_i^2$ as it traverses an angle of $\alpha = 2\pi$. For consistency, in the case of clockwise rotations (negative ω_i), we will consider the area swept out to also be negative.

Now, let $\psi'(t) = \psi(t + \mathrm{d}t)$. Then

$$\int_{\tau=0}^{t} \Im\langle\psi(\tau)|\psi'(\tau)\rangle = \int_{\tau=0}^{t} \Im \sum_i \bar{c}_i(\tau) c_i(\tau + \mathrm{d}\tau) \tag{22}$$
$$= \int \sum_i r_i^2 \Im\{e^{-i\theta_i(\tau)} e^{i[\theta_i(\tau) + \omega_i \mathrm{d}\tau]}\} \tag{23}$$
$$= \int \sum_i p_i \Im(1 + i\omega_i \mathrm{d}\tau) \tag{24}$$
$$= \int \sum_i p_i \mathrm{d}\alpha_i \tag{25}$$
$$= \int \mathrm{d}\alpha = \alpha(t) = A(t)[\psi(0)], \tag{26}$$

where the overbar denotes complex conjugation, $r_i = |c_i|$ as before, $\theta_i(\tau) = \arg(c_i(\tau))$, and α is now the weighted-average value of α_i.

Now, consider the *total* area $a(t)$ swept out by *all* coefficients c_i over time t. Note that $r_i^2 = |c_i|^2$ is also the probability p_i of basis state v_i, and so the *total* area swept out is always exactly half of the *average* angle $\alpha(t)$ of phase rotation (weighted by the state probability), or in other words, half of the expectation value of the $A(t)$ operator applied to the state $\psi(0)$. That is,

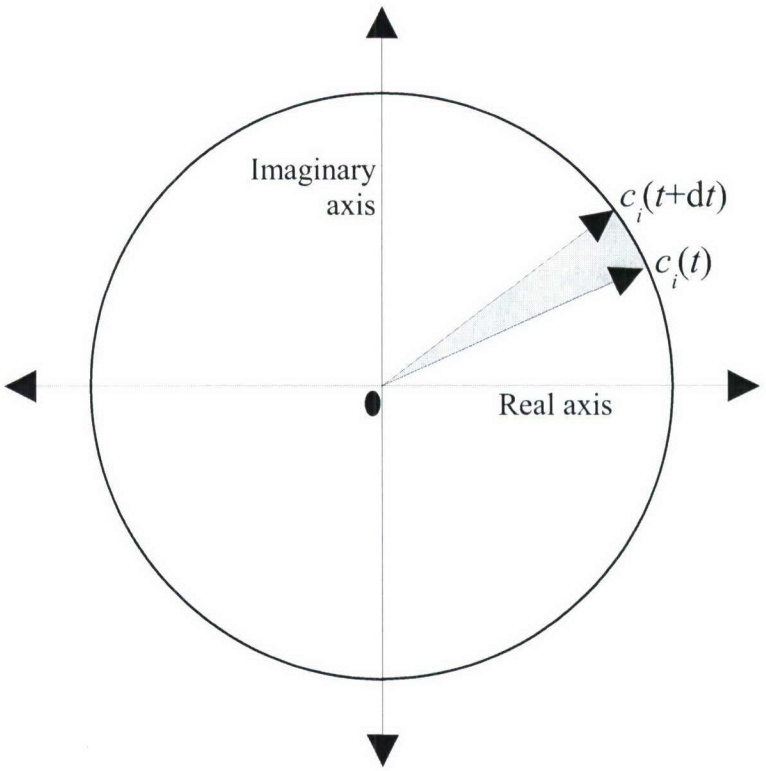


Fig. 3. In the energy eigenbasis, a complex coefficient c_i of a basis state sweeps out a small wedge-shaped area (shown exaggerated) in the complex plane over an infinitesimal time interval dt.

$$\begin{aligned} a(t) &= \sum_i \frac{1}{2}\omega_i t r_i^2 \\ &= \frac{1}{2}\sum_i p_i \alpha_i \\ &= \frac{1}{2}A(t)[\psi(0)] = \frac{1}{2}\alpha(t). \end{aligned} \tag{27}$$

Thus we have shown that for time-independent Hamiltonians, the expectation value of the action operator $A(t)$ applied to any initial state $\psi(0)$ is equal to the integral over the state trajectory of the inner product between infinitesimally adjacent states $\psi(t)$ and $\psi'(t)=\psi(t+dt)$ along the trajectory, as well as to the average phase angle α accumulated and to twice the complex-plane area a swept out by the state's coefficients, when the state is decomposed in the energy eigenbasis.

Of course, the inner product between two state vectors is a pure geometric quantity, and so is basis-independent. Therefore, the integral of $\Im\langle\psi|\psi'\rangle$ over the state trajectory does not depend at all on the (fixed) choice of basis under which states are decomposed into components. Likewise, the operator $A(t)$ itself is a geometric object not inherently associated with any particular basis. Therefore, the identity

$$\int_{\tau=0}^{t} \Im\langle\psi(\tau)|\psi'(\tau)\rangle = A(t)[\psi(0)] \tag{28}$$

that we proved above is a fundamental one whose truth does not rely on any particular basis or coordinate system.

However, it is perhaps somewhat less obvious that the average angle α of phase rotation and the complex-plane area a swept out by the state coefficients should also be basis-independent quantities, since their original definitions explicitly invoked a choice of basis (the energy basis). However, in the next section we will show that in fact, these quantities are basis-independent as well. Thus, all of the following identities still hold true, regardless of basis:

$$2a = \alpha = \int_{\tau=0}^{t} \Im\langle\psi|\psi'\rangle = A(t)[\psi(0)], \tag{29}$$

where a is the total complex-plane area swept out by the state coefficients in any fixed basis, $\alpha = \int \omega dt$ is the time-integral of the expected value ω of the angular velocity ω_i of the state coefficients in any fixed basis (not necessarily the same one), $\psi = \psi(\tau)$ is the state trajectory, with $\psi' = \psi(\tau + d\tau)$, $A(t)$ is the action operator as we defined in Eq. (10), and we are using our mean-value notation $A(t)[\psi(0)] = \langle\psi(0)|A(t)|\psi(0)\rangle$.

Our proposed measure of the amount of change undergone (and computational effort exerted) along a state trajectory $\psi(t)$ generated by a constant H will then just be the α value for that trajectory.

Later, in Sec. 8, we will show that the above identities also still hold even when $H(t)$ varies over time, and so our measure will generalize to that case as well.

7. GENERALIZING TO ARBITRARY BASES

The above discussion made use of a set of basis vectors $\{|v_i\rangle\}$ which were taken to be orthonormal eigenvectors of the (temporarily presumed constant) Hamiltonian operator H. Now, we will show that this particular choice of basis was in fact unnecessary, and that the same statements

concerning the relationship between the area swept out, the average phase angle accumulated, and the action $A(t)$ would remain true in any fixed (time-independent) basis.

At first, it may seem very non-obvious that the area swept out should still be exactly half of the action. Note that our previous arguments for this relied on the fact that in the energy basis $\{|v_i\rangle\}$, the coefficients c_i all rotate at uniform angular velocities ω_i in circles in the complex plane, while their individual magnitudes remain constant. In a different basis $|v_j\rangle$ (distinguished by using a different index symbol j), this will no longer be true. Each basis vector $|v_j\rangle$ in the new basis is in general some superposition of the $\{|v_i\rangle\}$, such as

$$|v_j\rangle = \sum_i u^i_j |v_i\rangle, \tag{30}$$

where the matrix $\mathbf{U} = [u^i_j]$ of complex coefficients (with the subscript j indexing rows, and the superscript i indexing columns) is, most generally, any unitary matrix. We can also write this equation in matrix-vector form as $\overrightarrow{|v_j\rangle} = \mathbf{U}\overrightarrow{|v_i\rangle}$, where the over-arrow here denotes that we are referring to the entire column-ordered sequence of basis vectors, $\overrightarrow{|v_i\rangle} = \begin{bmatrix} |v_1\rangle \\ \vdots \end{bmatrix}$. Of course, a general state vector ψ can equally well be expressed as a linear superposition of either set of basis vectors, that is,

$$|\psi\rangle = \sum_i c_i |v_i\rangle \tag{31}$$

$$|\psi\rangle = \sum_j c_j |v_j\rangle. \tag{32}$$

But now, we can substitute Eq. (30) into Eq. (32) and rearrange, as follows:

$$|\psi\rangle = \sum_{ij} c_j u^i_j |v_i\rangle = \sum_i \left(\sum_j c_j u^i_j \right) |v_i\rangle. \tag{33}$$

Now, since the $|v_i\rangle$ are linearly independent, the expansion of $|\psi\rangle$ in terms of them must be unique, so we can equate the coefficients on $|v_i\rangle$ in Eqs. (31) and (33) to get

$$\begin{aligned} c_i &= \sum_j u^i_j c_j \\ \overrightarrow{c_i} &= \mathbf{U}^{\mathrm{T}} \overrightarrow{c_j}, \end{aligned} \tag{34}$$

where T is matrix transpose. We can easily solve this equation for the c_j coefficients as follows:

$$\begin{aligned} \vec{c_i} &= \mathbf{U}^{\mathrm{T}} \vec{c_j} \\ (\mathbf{U}^{\mathrm{T}})^{-1} \vec{c_i} &= \vec{c_j} \\ \bar{\mathbf{U}} \vec{c_i} &= \vec{c_j} \\ c_j &= \sum_i \bar{u}^i_j c_i . \end{aligned} \tag{35}$$

In other words, each complex coefficient in the new basis is just a particular linear combination of what the various complex coefficients were in the old basis.

If the coefficients c_i in the old energy basis are describing perfect circles around the complex origin at a variety of radii and angular velocities, there is no guarantee that the coefficients c_j in the new basis will still be describing circular paths centered on the origin, although their paths will of course still be continuous and smooth if the original c_i trajectories were. In general, the c_j will follow complicated looping trajectories in the complex plane, generated as if by Ptolemaic planetary epicycles, i.e., as a sum of circularly rotating vectors. A given c_j will in general return to its initial location in the complex plane only when its components c_i that have non-zero values of u^i_j all simultaneously return to their initial locations exactly, which might even take infinitely long, if the corresponding ω_i values were relatively irrational.

Anyhow, the important point for our present purposes is that the c_js do not, in general, maintain a constant magnitude (distance from the origin), and so the area swept out by the c_j over a given time is no longer just a section of a circle, which was very easy to analyze. Instead, while c_j's phase angle θ_j is rotating, simultaneously its magnitude r_j may also be growing or shrinking. Figure 4 illustrates the situation.

To clarify what we mean by the phase angle $\theta_j(t)$ a bit more carefully, let us use $\mathrm{d}\alpha_j(t) \approx 0$ to denote the infinitesimal increment of phase angle from times t to $t + \mathrm{d}t$ such that

$$\mathrm{d}\alpha_j \equiv \arg(c'_j) - \arg(c_j) \pmod{2\pi}, \tag{36}$$

so that $\mathrm{d}\alpha_j$ remains infinitesimal even when c_j crosses a branch cut of the Arg() function. Then, let $\alpha_j(t)$ be the total accumulated phase angle over time t, that is, the integral of $\mathrm{d}\alpha_j$ over time,

$$\alpha_j(t) = \int_{\tau=0}^{t} \mathrm{d}\alpha_j \tag{37}$$

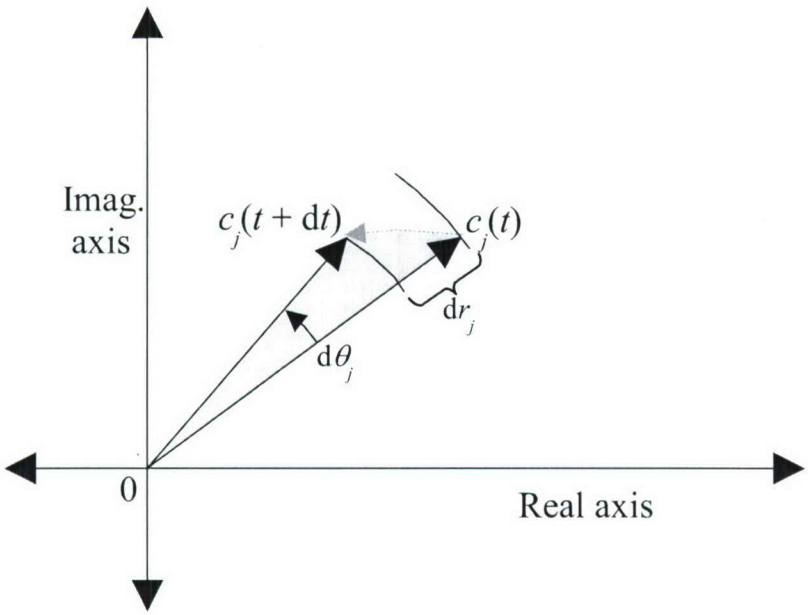


Fig. 4. Area swept out (exaggerated) by a coefficient c_j (in a basis other than the energy eigenbasis) over an infinitesimal time interval $\mathrm{d}t$. Note that both its phase and its magnitude change, in general.

so that $\alpha_j(0) = 0$. Now, just let $\theta_j(t) = \mathrm{Arg}[c_j(0)] + \alpha_j(t)$. Thus also $\mathrm{d}\theta_j = \mathrm{d}\alpha_j$.

What, now, is the area swept out in our new basis? First, notice that in the infinitesimal limit, it is exactly half of the area of the parallelogram that is spanned on two adjacent sides by $c_j = c_j(t)$ and $c'_j = c_j(t+\mathrm{d}t)$, considered as vectors in the complex plane. See Fig. 5.

The parallelogram area, itself, is $\mathrm{d}a_j = r_j r'_j \sin(\mathrm{d}\theta_j)$, where r_j and r'_j are the magnitudes of the old and new coefficients, respectively. However, note that the area $\mathrm{d}a_j$ of this parallelogram is also the signed magnitude

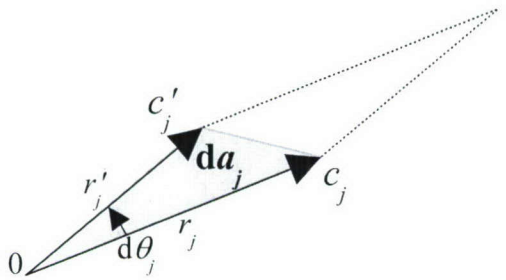


Fig. 5. The infinitesimal area $\mathrm{d}a_j$ swept out approaches one-half of the parallelogram area $r_j r'_j \sin \mathrm{d}\theta_j$.

of the scalar "cross product" $c_j \times c'_j$ between the coefficients, considered as vectors in the complex plane. (The traditional cross product, defined in three dimensions, would be a vector perpendicular to the complex plane having this value $\mathrm{d}a_j$ as its length.) There is a nice identity[20] connecting the scalar cross product and dot product with the conjugate multiplication of complex numbers, namely:

$$\bar{c}d = c \cdot d + \mathrm{i}(c \times d), \tag{38}$$

where $\bar{c}$ means the complex conjugate of c, and $c \cdot d$ denotes the real scalar "dot product" between c and d considered as vectors, namely $|c||d|\cos[\arg(d) - \arg(c)]$, and $c \times d$ denotes the real scalar "cross product" previously mentioned, namely $|c||d|\sin[\arg(d) - \arg(c)]$.

Applying this identity to our situation, we can see that the area swept out, since it is half the cross product, is half of the imaginary part of the conjugate product $\bar{c}_j c'_j$ between the old and new coefficients, and also to half of $\sin(\mathrm{d}\alpha_j) = \mathrm{d}\alpha_j$;

$$\mathrm{d}a_j = \frac{1}{2}\mathrm{d}\alpha_j = \frac{1}{2}\Im(\bar{c}_j c'_j). \tag{39}$$

Now, this is just the area swept out by a single component c_j. To find the total area $\mathrm{d}a$ swept out by all coefficients, we merely sum over components:

$$\begin{aligned} \mathrm{d}a &= \frac{1}{2}\sum_j \Im(\bar{c}_j c'_j) = \frac{1}{2}\Im\sum_j \bar{c}_j c'_j \\ &= \frac{1}{2}\Im\langle\psi|\psi'\rangle = \frac{1}{2}\mathrm{d}\alpha \end{aligned} \tag{40}$$

In other words, *just like in the energy basis*, in an arbitrary basis, it is still true that the infinitesimal increment $\mathrm{d}a$ in the area swept out by the coefficients is exactly one-half of $\Im\langle\psi|\psi'\rangle$, the imaginary component of the inner product between infinitesimally adjacent vectors $\psi = \psi(t)$ and $\psi' = \psi(t + \mathrm{d}t)$ along the trajectory, and further that this is equal to half of $\mathrm{d}\alpha = \mathrm{d}\theta$, the average increment of the continuously varying phase angles $\theta_j(t)$ of the coefficients.

Now, we saw earlier that $\Im\langle\psi|\psi'\rangle$ is also equal to the expectation value $A'[\psi] = \langle\psi|A'|\psi\rangle$ of the infinitesimal action operator $A' = H\mathrm{d}t$ applied to the state ψ, for any state ψ. So in connection with the result (40) that we just obtained, this means that $A'[\psi]$ gives *exactly* the average phase angle accumulation $\mathrm{d}\alpha$ of the coefficients c_j of ψ in *any* basis, and twice the complex-plane area $\mathrm{d}a$ swept out by those coefficients. We can

thus think of A' as being the operator representation of a fundamental, basis-independent concept of "average angle accumulated" or "total area swept out" over infinitesimal intervals.

8. GENERALIZING TO TIME-DEPENDENT HAMILTONIANS

In the previous section, we established the basis-independence of the identities $2\mathrm{d}a = \mathrm{d}\alpha = \Im\langle\psi|\psi'\rangle = \omega\,\mathrm{d}t = A'[\psi] = \langle\psi|H\mathrm{d}t|\psi\rangle$ for infinitesimal changes of the state vector ($\psi \to \psi'$) along its trajectory over infinitesimal time intervals $\mathrm{d}t$, under any *constant* Hamiltonian H.

But, as long as the Hamiltonian $H(t)$ only changes in continuous fashion, it can always be considered essentially "constant" throughout any infinitesimal interval $\mathrm{d}t$, even if it is varying over non-infinitesimal time-scales. Therefore, the above identities will still hold true instantaneously even for a time-dependent Hamiltonian $H(t)$, which is what we originally started out our discussion with. Thus, when we integrate the above Eq. (40) over time, it remains true that:

$$2a = \alpha = \int_{t=t_1}^{t_2} \Im\langle\psi(t)|\psi(t+\mathrm{d}t)\rangle \tag{41}$$

$$= \int_{t=t_1}^{t_2} \omega(t)\mathrm{d}t \tag{42}$$

$$= \int_{t=t_1}^{t_2} \langle\psi(t)|H(t)|\psi(t)\rangle\mathrm{d}t \tag{43}$$

$$= \int_{t=t_1}^{t_2} A'(t)[\psi(t)]. \tag{44}$$

In words, this says that for any initial state ψ, we have that $2a$ (twice the complex-plane area swept out by the coefficients of ψ, in any basis) is equal to α, the average phase angle swept out by the state coefficients, as well as to (41) the integral along the trajectory $\psi(t)$ of the imaginary component of the dot product between neighboring vectors along the trajectory, and also to (42) the integral of the average phase velocity of the coefficients, weighted by the instantaneous basis state probabilities $p_i(t) = r_i(t)^2$, which is (43) the time-integral of the instantaneous Hamiltonian energy $E(t) = H(t)[\psi(t)]$ of the instantaneous state $\psi(t)$, which (finally) is (44) the integral of the infinitesimal actions $\mathrm{d}\alpha(t) = \langle\psi(t)|A'(t)|\psi(t)\rangle$ on the instantaneous states $\psi(t)$.

The natural next question to ask is, given that $A'[\psi] = \mathrm{d}\alpha$ remains true over infinitesimal intervals $\mathrm{d}t$ in the general time-dependent case,

and given that cumulatively, $A(t)[\psi(0)]=\alpha$ in the time-independent case ($H(t)=H=$const.), does this cumulative relation still hold true in the general time-dependent case? That is, for $A(t)$ (as defined in Eq. (10)) is it still true that

$$A(t)[\psi(0)]=\alpha \tag{45}$$

even if the phase angle α was accumulated under the influence of a varying $H(t)$?

If this Eq. (45) is universally correct, then we will have a very nice, simple interpretation for the general action operator $A(t)$ even in the case of a time-dependent $H(t)$, namely that, when applied to any initial state $\psi(0)$, it simply gives the angular length α of the trajectory that will be traversed by that state, a quantity which obeys all of the identities (41)–(44).

Actually it seems that this is true, and the proof is quite elegant. First, from Eq. (17) and the boundary condition $U(0)=1$, fix $U=U(t)$, the overall unitary transform operating between times 0 and t that is implied by the values of the time-dependent Hamiltonian $H(\tau)$ for all $0\leq\tau\leq t$. Fix then also $A=A(t)$ by using Eq. (13) and the associated discussion, using the continuity requirement on $A(\tau)$ and the requirement that $A(0)=0$.

Now, consider any eigenvector $|\phi_i\rangle$ of U, which is a state that undergoes a cyclic evolution (in the projective Hilbert space) under $H(\tau)$ or any other process (Hamiltonian trajectory) that implements U, since $U|\phi\rangle=\mu_i|\phi_i\rangle$, with μ_i being the associated unit-modulus eigenvalue. Of course, $|\phi_i\rangle$ is then also an eigenvector of A, with an eigenvalue α_i such that $A|\phi_i\rangle=\alpha_i|\phi_i\rangle$ and $\mu_i=e^{i\alpha_i}$.

To see that this α_i must indeed be the same as the total phase angle α accumulated by $|\phi_i\rangle$ as defined in e.g. Eq. (44), consider that once the overall operator A has been determined, we can simply divide it by t to find an alternative *time-independent* $H_c=A/t$ that would also generate the very same action operator A and the same unitary U when applied over the same time interval t. From the discussion in Sec. 6, its easy to see that the value of α is then indeed exactly the phase angle accumulated from the initial state $|\phi_i\rangle$ when implementing A via this (alternative) time-independent H_c.

Now, does *every* Hamiltonian trajectory that implements A (including our original time-dependent $H(\tau)$) involve the same total accumulation α of phase angle? We can see that it must, because any trajectory $H(\tau)$ can, it seems, be continuously deformed into the constant trajectory $H_c(\tau)=H_c$ while maintaining the same overall A (and thus U) throughout the deformation process. At no point during this continuous deformation process can

the total phase α that is accumulated ever change, since, to produce the same U, the total phase α must always remain congruent to $\alpha_i \pmod{2\pi}$, and it would be impossible for the total phase accumulated to jump by a multiple of 2π at any point during any continuous deformation of the trajectory.

To see that this is true, recall from Eq. (13) and the associated discussion that any continuous $A(\tau)$ can be characterized by a continuously varying eigenbasis $\{|u_i(\tau)\rangle\}$ of $U(\tau)$ (with a sort of k-dimensional continuous gauge freedom, where k is the Hilbert space dimension), and by implied integer parameters $n_i(\tau)$ that select which of the logarithm values must be used at each time point τ. As we continuously deform the Hamiltonian trajectory $H(\tau)$ as well as the eigenbases $\{|u_i(\tau)\rangle\}$ (and thus the gauges of the associated eigenvalues $u_i(\tau)$), the set of time points τ at which the $n_i(\tau)$ values change also changes continuously. Nowhere during this continuous, local process can the total angle α accumulated along the trajectory possibly change discontinuously by a multiple of 2π.

Thus, our arbitrary time-dependent $H(\tau)$ takes the eigenstate $|\phi_i\rangle$ through the same total angle α as would the constant H_c for which we already know that $\langle\phi_i|A|\phi_i\rangle = \alpha$.

The above discussion establishes that (regardless of the dynamics $H(t)$) the A operator that we derive from it always gives the correct accumulated angle α for all eigenstates ϕ_i of A; therefore it is also correct for arbitrary initial superposition states $\psi(0)$ (and for mixed states as well).

For a final interesting observation, let $\alpha(\psi(0), t)$ denote the angle α accumulated from the initial state $|\psi(0)\rangle$ over time t, and note that since

$$\langle\psi(0)|A(t)|\psi(0)\rangle = \alpha(\psi(0), t) \tag{46}$$

for all initial $\psi(0)$, the time-derivative of the operator $A(t)$ must satisfy

$$\langle\psi(0)|\frac{\mathrm{d}}{\mathrm{d}t}A(t)|\psi(0)\rangle = \frac{\partial}{\partial t}\alpha(\psi(0), t). \tag{47}$$

Recall meanwhile that $\mathrm{d}\alpha(t)$ is given by applying $A'(t) = H(t)\mathrm{d}t$ to the state $\psi(t)$; i.e., $\mathrm{d}\alpha(t) = A'(t)[\psi(t)]$. Of course, $\psi(t) = U(t)\psi(0)$, so we have that

$$\langle\psi(0)|\frac{\mathrm{d}A}{\mathrm{d}t}(t)|\psi(0)\rangle = \frac{A'(t)}{\mathrm{d}t}[U(t)\psi(0)] \tag{48}$$

$$= \langle\psi(0)|U^{\dagger}(t)H(t)U(t)|\psi(0)\rangle. \tag{49}$$

and thus

$$\frac{\mathrm{d}A}{\mathrm{d}t}(t) = U^{\dagger}(t)H(t)U(t)$$

$$= \mathrm{e}^{-\mathrm{i}A(t)}H(t)\mathrm{e}^{\mathrm{i}A(t)}. \tag{50}$$

Now, note that applying the time-dependent operator form (6) of the Schrödinger equation to $U(t) = e^{iA(t)}$, we get

$$\begin{aligned} \frac{d}{dt} e^{iA(t)} &= iH(t)e^{iA(t)} \\ &= ie^{iA(t)}e^{-iA(t)}H(t)e^{iA(t)} \\ &= e^{iA(t)}\frac{d}{dt}[iA(t)], \end{aligned} \tag{51}$$

where we have used (50) in the last step. In other words, the ordinary rule $de^f = e^f df$ for the differential of an exponential of a function f actually turns out to be true when $f = iA(t)$, despite the fact that the Hamiltonian may be time-dependent and that $A(t)$ doesn't necessarily even commute with its time-derivative! This is due to the special way in which we defined our $A(t)$ function, and would not be true for more general time-dependent operators.

9. DISCUSSION OF EFFORT

Although a choice of a particular cumulative action operator A still gives us freedom to choose any number of different Hamiltonian trajectories $H(\tau)$ for implementing it, over various total amounts of time t, we have seen above that all such trajectories are equivalent in terms of the total amount α of phase angle that is accumulated starting from any fixed initial state $|\psi(0)\rangle$.

As hinted previously, we might even consider the quantity α (or, more properly, its absolute value) to be a reasonable definition of the geometric *length* of the path that a normalized state vector $|\psi(t)\rangle$ describes as it moves along any continuous path (parameterized by any real variable t) along the unit sphere in Hilbert space, since (note) α depends only on the shape of the state trajectory itself, and not on any other properties of the Hamiltonian trajectory, such as the energy of other orthogonal states.

As a result, an intrinsic metric on the normalized Hilbert space is provided by the distance function

$$d(|\psi_1\rangle, |\psi_2\rangle) :\equiv \min |\alpha|, \tag{52}$$

where α is the accumulated phase angle along a given trajectory, and the minimum is taken over all normalized, continuous paths from $|\psi_1\rangle$ to $|\psi_2\rangle$, or a subset of such that is deemed available. The absolute-value operator is required in order to obtain a proper (positive) metric, since trajectories with unboundedly negative values of α could exist if we allow states

to have negative energy. Paths having the minimum absolute α between a given pair of states can be considered to be (sections of) geodesics on the normalized Hilbert space.

In Ref. 21, Wootters introduced a statistically motivated distance metric between quantum states which he called "statistical distance," and showed that it was identical to the ordinary Hilbert-space distance function $d(\psi_1, \psi_2) = \arccos|\langle\psi_1|\psi_2\rangle|$. It turns out that our distance function d above is in fact exactly the same as this also, if all Hilbert-space trajectories are considered. However, if the space of allowed trajectories is restricted (for example, if the Hamiltonians are forced to be local) then a different distance measure results. In Wootters' metric, the distance between any two distinguishable states (e.g., two different randomly chosen computational basis states) is only $\arccos 0 = \pi/2$, whereas if we define distance by minimizing over allowed trajectories, we could obtain a much greater figure.

Later, we will see that our distance measure will also allow us to derive a natural metric on unitary operations, telling us the "distance" between two unitaries, as measured by the difficulty of getting from one to the other, in terms of the minimum distance traversed by worst-case states.

Anyway, noting that this measure α of trajectory length which we have explored above is stable with respect to changes of basis, that there are multiple simple ways of defining it, and that it connects strongly with fundamental physical concepts such as action and energy, as well as with primitive geometric concepts such as angles and areas, and that it forms a natural metric on the Hilbert space, all of these facts together motivate us to propose this measure as being the most natural and genuine measure of the total "amount of change" that is undergone by a physical quantum state vector $|\psi(t)\rangle$ as it changes dynamically under a (possibly varying) physical influence $H(t)$.

Insofar as we can consider *all* dynamical evolution and change to be forms of "computation," where this word is construed in a very general sense, we can also accept this measure as being an appropriate measure of the *amount of computational effort exerted* by the system as it undergoes the given trajectory.

Thus, from here on, rather than calling our quantity "action" (which would lead to confusion with the action of the Lagrangian), or "accumulated phase angle" (which is awkward) we will refer to our quantity as simply the *effort* when we wish to be concise, and abbreviate it with the symbol $\mathcal{F}$. That is,

$$\mathcal{F}_{t_1 \to t_2}[\psi(t)] :\equiv \int_{t=t_1}^{t_2} \Im\langle\psi(t)|\psi'(t)\rangle \tag{53}$$

is a real-valued functional of a state vector trajectory $\psi(t)$ taken between two times t_1 and t_2. Note that the value of $\mathcal{F}$ depends only on the shape of the path. It is independent of the absolute time, the speed at which the trajectory is traversed, and on various other details of the Hamiltonian that generates the trajectory (such as its eigenvalues for eigenstates that are not components of ψ); in general, many different Hamiltonian evolutions can generate the same path, which will always have the same total effort. So, in the above equation, we can consider $\psi(t)$ to just be a parameterized curve where t is now just any *arbitrary* real-valued parameter, not necessarily even corresponding to physical time. In other words, the *effort* quantity does not depend on the precise system of coordinates that is used for measuring the passage of time, but rather only on a pure geometric object, namely the path taken through Hilbert space.

Note that to say that the path length corresponds to computational *effort* is not to imply that all of the physical computation that is occurring in the given system is necessarily being harnessed and applied by humans to meet our calculational needs, only that this is the total amount of raw computational work that is occurring "in nature." The choice of the word "*effort*" is intended to evoke the commonsense realization that effort may be wasted, i.e., not used for anything useful.

Note also that the action operator A (as we have defined it) gives a concise yet particularly comprehensive characterization of a given computational process, in the sense that it determines not only the overall unitary operation $U = e^{iA}$ that will be performed, but also the amount of effort that will be expended in getting to the final result from any given initial state.

The primary caveat to the above conception of computational effort seems to be that the quantity $\mathcal{F}$ (together with the rate of phase rotation, and the path length in Hilbert space) is dependent on where we choose to draw our zero of energy. As is well known, absolute energies are only physically defined up to an additive constant, and so the total Hamiltonian action or effort is only well defined up to this constant multiplied by the elapsed time t.

A natural and widely used convention is to define the least eigenvalue of the Hamiltonian (the "ground state" energy) to be the zero of energy. In a similar fashion, we can choose to additively shift the Hamiltonian so that the least eigenvalue of the cumulative action operator $A(t)$ is taken to represent zero effort. (Note that this approach can even be used when the Hamiltonian itself is time-dependent.)

However, this choice is by no means mandated mathematically, and in fact, in certain pathological cases (such as an infinite-dimensional or time-dependent Hamiltonian with unboundedly negative eigenvalues),

there might not even be *any* minimum eigenvalue for the resulting action operator over a given interval. One needs to keep these caveats in the back of one's mind, although they seemingly end up not very much affecting the potential practical applications of this concept, which we will address in a later section.

Another reason that we might not want to consider the ground state energy to always be zero is if the ground state energy varies, especially if it includes energy that had to be explicitly transferred into the system from some other external subsystem. Thus, energy that is present in a given system, even if that system is in its ground state, may still represent energy that was transferred from elsewhere and isn't being used for other purposes; i.e., it may represent "wasted" computational effort, and we may wish to count it as such, rather than just counting it as zero effort.

Another possible convention would be to count a system's energy as being its total (gravitating) mass–energy, or rest mass–energy, if we want it to be independent of the observer's velocity. One might think this choice is a somewhat less arbitrary than the ground state convention, since mass is a physical observable, but unfortunately, in general relativity, the contribution to the total mass–energy of a local system that is due to its gravitational self-energy isn't actually independent of the coordinate system that is used (Ref. 22, p. 62). However, this caveat is usually only important in extreme systems such as neutron stars and black holes, where the gravitational self-energy contributes significantly to the system's total mass.

In any case, for now, we propose to just make a "gentlepersons' agreement" that we will always make sure that the energy eigenvalues of the systems that we consider are always shifted so as to be positive, so that the total effort is always positive, and we don't have to worry about what would be the meaning of a negative "amount of computational effort." Unfortunately, this strategy rules out considering certain classes of systems, such as bottomless potential wells, or the infinite Dirac sea of negative-energy fermion states. But resolving this issue will have to wait for future work.

10. MORE ABSTRACT SCENARIOS

In the above, we have specified a well-defined (at least, up to an additive constant) positive, real-valued measure $\mathcal{F}$ of the amount of computational effort represented by any trajectory of a state vector in Hilbert space.

This raises the question of whether we can assign a measure of computational effort to other physical situations that may be less completely specified. For example, we may be given a cumulative action operator A,

but not know the detailed Hamiltonian trajectory $H(t)|_{t=t_1}^{t_2}$ that generated it, and we may be given only a set V of possible initial states (rather than a single definite state), or we may have a probability distribution or density function $p: V \to [0, 1]$ over initial states. In such more abstract situations, can we still meaningfully define the amount of computational effort exerted by the system as it undergoes the evolution specified by its Hamiltonian over a given time interval?

Of course we can. Given a cumulative action operator A and given any specific state $\psi = \psi(t_1)$ at the initial time t_1, the value of $\mathcal{F}_{t_1 \to t_2}[\psi(t)]$ is independent of the details of the Hamiltonian trajectory $H(t)$ and is given simply by

$$\mathcal{F}_A(\psi) :\equiv A[\psi] = \langle \psi | A | \psi \rangle, \tag{54}$$

which can be called *the effort undergone by ψ under A*.

We can therefore also naturally express the *average or expected effort over V exerted by the action operator A* as:

$$\widehat{\mathcal{F}}_V(A) = \mathrm{Ex}_V[\mathcal{F}_A] = \sum_{\psi \in V} p(\psi)\mathcal{F}_A(\psi) = \langle A \rangle = \mathrm{Tr}(\rho A), \tag{55}$$

where the density operator ρ describing the initial mixed state is constructed from the probability distribution over pure states ψ in the usual fashion, that is, with $\rho = \sum_{\psi \in V} p(\psi)|\psi\rangle\langle\psi|$. If no probability distribution p has been provided, we can use a uniform distribution over some natural measure on the set V.

This then gives us a workable definition of the mean effort exerted by a system over time under a given Hamiltonian, even when the initial state is not exactly known.

In some situations, we might also be particularly interested in the *maximum* effort over the set V of possible initial states. For example, suppose we are preparing the initial state of the system, and we want to initialize the system in such a way that it will exert the maximum effort possible. Given A and maximizing over V, we define the *maximum effort exerted by A over V* as

$$\mathcal{F}_V^+(A) :\equiv \max_{\psi \in V} \mathcal{F}_A(\psi). \tag{56}$$

This can be considered to be a measure of the *potential* computational "strength" of the given action operator A, expressing that any Hamiltonian $H(t)$ that implements A over some arbitrary interval $t_1 \to t_2$ *could* exert an amount $\mathcal{F}_V^+(A)$ of computational effort over that same interval, given a suitable initial state. Insofar as the actual state that we end up getting

might be the one that undergoes the maximal amount of effort, we can say that a system with an unknown or unspecified state is, at least, exerting this much "potential" computational effort.

Even if the actual state turns out *not* to be the maximal-action one, the system could still be thought of as having "done the work" of determining that the actual state is *not* the one that should have transitioned through the given maximum Hilbert-space distance. This particular thought should really be credited to Seth Lloyd, who pointed out to me in personal discussions, as an analogy, that an ordinary Boolean gate operation can still be thought of as doing computational work even if the output bit that it is applied to is not actually changed; namely, it is doing the work of determining *that* the bit should not change.

Similarly to how we defined the maximum effort, we can likewise define the *minimum effort of A over V* as $\mathcal{F}_V^-(A) :\equiv \min_{\psi \in V} \mathcal{F}_A(\psi)$, although we should keep in mind that if the ground state of the action operator A is an available initial state in V, and if we use the convention that the ground state action is defined to be zero, then $\mathcal{F}_V^-(A)$ will always be 0, and so will not be very useful.

11. DIFFICULTY OF PERFORMING AN OPERATION

Suppose now that we are given *no* information about the situation to be analyzed except for a unitary operator U on the Hilbert space $\mathcal{H}$, and we want to address the following question: How much computational effort, at minimum, is required to physically implement U? By "implement" we mean that U is the time evolution operator that ends up being generated by the dynamics over some interval, according to $U = e^{iA}$ for some action operator A. We can call this minimum required effort the *difficulty* $\mathcal{D}$ of implementing the unitary operator U. Our framework gives us a natural way to formalize this notion.

Assuming we have some freedom of choice in the design of the system, then among the set $\mathcal{A}$ of all Hermitian operators A on $\mathcal{H}$, or among at least a set $\aleph \subseteq \mathcal{A}$ of *available* or implementable action operators, we might want to choose the operator A that generates U that has the *smallest* value of the maximum or worst-case effort $\mathcal{F}_V^+(A)$ over the set V of possible initial state vectors. This A can be considered to be the "best" action operator for generating the given unitary U, in the sense that the length of the longest trajectory that would be undergone by any possible state vector $\psi \in V$ is minimized. This strategy is analogous to what we do in traditional algorithm design, where we usually choose the algorithm that has the minimum time complexity on worst-case input data. In our

case, A can be considered to abstractly represent the algorithm selected, while the initial vector ψ represents the input data. Rather than time complexity, we focus on effort or Hamiltonian action, since (as we will see) this translates directly to time when a given supply of energy is available to be invested in the system.

In some situations, it may be preferred to choose A so as to minimize the *expected* effort rather than the worst-case effort, for example, if we want to minimize the total effort exerted over an arbitrarily large set of computations with randomly chosen input states selected from some distribution.

We can thus define the maximum ($\mathcal{D}^{+}_{\aleph,V}$) and expected ($\widehat{\mathcal{D}}_{\aleph,V}$) difficulty of a desired unitary transform U under the available action set $\aleph$ and initial-state set V as follows:

$$\begin{aligned}\mathcal{D}^{+}_{\aleph,V}(U) &:\equiv \min_{A\in\aleph} \mathcal{F}^{+}_{V}(A) \\ &= \min_{A\in\aleph} \max_{\psi\in V} \mathcal{F}_{A}(\psi) \qquad (57)\\ \widehat{\mathcal{D}}_{\aleph,V}(U) &:\equiv \min_{A\in\aleph} \widehat{\mathcal{F}}_{V}(A) \\ &= \min_{A\in\aleph} \sum_{\psi\in V} p(\psi)\mathcal{F}_{A}(\psi) \qquad (58)\end{aligned}$$

Note that in all cases we still want to minimize over the available action operators $A \in \aleph$, because there is usually no physical reason why indefinitely large action operators (which waste arbitrarily large amounts of effort) could not be constructed to implement a given unitary; thus, maximizing over action operators would thus always give ∞ and would not be meaningful.

A remark about the set $\aleph$ of available action operators. Typically it would be constrained by what constitutes an "available" dynamics that we are free to choose within a given theoretical, experimental, or manufacturing context. For example, $\aleph$ might reasonably be constrained to include only those action operators that are obtainable from time-dependent Hamiltonians $H(t)$ which are themselves constructed by summing over local interaction terms between neighboring subsystems, or by integrating a Hamiltonian density function that includes only local terms on a field over some topological space, e.g., to reflect the local structure of spacetime in a quantum field theory picture. Or, we might constrain ourselves to action operators that are obtainable from time-independent Hamiltonians only, e.g., if we are designing a self-contained (closed) quantum system. Finally, practical considerations may severely constrain the space of Hamiltonians to ones that can be readily constructed in devices

that can be built using a specific manufacturing process, although we should note that if scalable universal quantum computers can be built, then any desired local Hamiltonian could be straightforwardly emulated on these machines.

As a brief aside, it is also interesting to note that a given difficulty function $\mathcal{D}(U)$ (either the worst-case or average-case version, and whatever $\aleph$ and V are) also induces an intrinsic metric on the space of unitaries of a given rank; we can define a suitable distance function between unitaries by

$$d(U_1, U_2) = \mathcal{D}(U_2 U_1^\dagger) \tag{59}$$

that is, the distance between U_1 and U_2 in this metric is just the difficulty of performing the relative unitary $U_{1\to 2} :\equiv U_2 U_1^\dagger$ that is equivalent to undoing U_1 (using $U_1^\dagger = U_1^{-1}$) and then doing U_2. A unitary trajectory for implementing $U_{1\to 2}$ that actually minimizes the effort will then form, when right-multiplied by U_1, a (section of a) geodesic in the space of unitaries passing between the unitaries U_1 and U_2 (since $U_{1\to 2} U_1 = U_2$). Of course, in general, the shortest unitary trajectory for implementing $U_{1\to 2}$ will *not* actually work by doing $U_1^\dagger$ followed by U_2; for example, if U_1 and U_2 have high difficulty but are very close together, then the shortest unitary trajectory between them will be much more direct than this.

Now, given our notion of the computational difficulty of a given unitary U, we can now reinterpret previous results (such as[5,16]) regarding "quantum speed limits" or minimum times to implement various specific unitary transforms of interest, or classes of transforms, given states of specified average energy above the ground state, as follows: These analyses are implicitly specifying an $\aleph$ (usually, just all Hermitian operators) and a V (usually, just the entire Hilbert space), and showing that the worst-case difficulty $\mathcal{D}^+(U)$ for the transform U has a specific value (or lower bound), assuming the presence of a time-independent Hamiltonian where the ground state energy is usually set to 0. In other words, such analyses show that a certain minimum worst-case effort or Hamiltonian action is required to implement the particular U in question.

As an example, Margolus and Levitin's result[5] can be interpreted as telling us that any U that rotates some state ψ to an orthogonal state has a worst-case difficulty of $\mathcal{D}^+(U) \geq h/4$, since their result shows that any state of energy E takes time at least $h/4E$ (no matter what the Hamiltonian) to accumulate the action needed to take it to an orthogonal state; thus the Hamiltonian action $A = Et$ that is required to carry out such a transition is at least $h/4$.

Another result in Ref. 5 implies that if there is a ψ such that ($|\psi\rangle$, $U|\psi\rangle$, $U^2|\psi\rangle$, ..., $U^{N-1}|\psi\rangle$, $U^N|\psi\rangle = |\psi\rangle$) comprises a cycle of N states, with each orthogonal to the preceding and succeeding states in the cycle, then $\mathcal{D}^+(U) \geq \frac{h}{2}\frac{N-1}{N}$, even if we are given complete freedom in constructing the Hamiltonian, aside from a requirement that it be time-independent. For $N = 2$, this expression reduces to $h/4$, while for $N \to \infty$, it goes to $h/2$. Thus, any physical computation that proceeds autonomously though an unbounded sequence of distinct states must exert at least $h/2$ effort per state transition.

Notice that the Margolus–Levitin theorem is, strictly speaking, only giving us a *lower bound* on the worst-case difficulty, since it is considering only a particular state ψ of interest (namely, one that actually undergoes a transition to an orthogonal state), rather than finding the worst-case potential effort to perform the corresponding U, maximized over all possible initial ψ in the Hilbert space. Later, we will see that the actual worst-case effort for an orthogonalizing transformation is actually $h/2 = \pi$ even in the $N = 2$ case, and possibly even higher in cases that go through more states.

We anticipate that, armed our definitions, it would be a highly useful and worthwhile exercise to systematically go through a variety of the quantum unitary transforms that have already been identified in quantum computing as comprising useful "quantum logic gate" operations, and quantify their worst-case and average difficulty, according to the above definitions, under various physically realistic sets of constraints. This would directly tell us how much physical Hamiltonian action is required to carry out those operations (given a best-case Hamiltonian implementation, while operating on a worst-case or average-case input state). We can likewise do the same for classical reversible Boolean logic operations embedded within unitary operations, as well as classical irreversible Boolean logic operations embedded within classical reversible operations, with ancilla bits used as needed for carrying away garbage information to be discarded.

Such an investigation will, for the first time, give us a natural and physically well-founded measure of the physical complexity of logic operations, in terms of Hamiltonian action. This in turn would directly tell us the minimum physical time to perform these operations within any physical system or subsystem using a set of states having a given maximum energy about the ground state, given the known or prespecified constraints on the system's initial state and its available Hamiltonian dynamics. This new quantification of computational complexity may also allow us to derive lower bounds on the number of quantum gates of a given type that would be required to implement a given larger transformation in

terms of smaller ones, and possibly to show that certain constructions of larger gates out of smaller ones are optimal.

In subsequent subsections, we begin carrying out the above-described line of research, with some initial investigations of the difficulty of various simple operations in situations where the available dynamics is relatively unconstrained, which is the easiest case to analyze.

12. SPECIFIC OPERATIONS

In this section, we explore the difficulty (according to our previous definitions) of a variety of important quantum and classical logic operations.

We will begin by considering some educated guesses about the difficulty of various unitaries. For each unitary U we are to imagine implementing it via a particular transformation trajectory $U'(t)$ (and Hamiltonian $H(t)$ such that $U'(t) = e^{iH(t)dt}$) that is as "direct" as possible, in the sense of minimizing the Hilbert-space distance through which worst-case states are transported. Intuition tells us that these minimal trajectories are expected to follow geodesics in the space of unitaries, as per the metric we defined earlier; in other words, they should be "straight-line" paths, so to speak, that get us to the desired unitary as directly as possible.

12.1. General Two-dimensional Unitaries

Let us begin by considering U_2, the space of unitary transformations on Hilbert spaces of dimensionality 2. In quantum computing, these correspond to single-qubit quantum logic gates. As is well known (e.g., see Ref. 23, Eq. 4.9), any such U can be decomposed as

$$U = e^{i\alpha} R_{\hat{n}}(\theta), \tag{60}$$

where $\hat{n} = (n_x, n_y, n_z)$ is a real 3D unit vector and $R_{\hat{n}}(\theta)$ is a Bloch-sphere rotation about this vector by an angle of θ, that is,

$$R_{\hat{n}}(\theta) = e^{i(\theta/2)(\hat{n}\cdot\boldsymbol{\sigma})}, \tag{61}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{62}$$

Let us now consider breaking down U into its multiplicative factors $e^{i\alpha}$ and $R_{\hat{n}}(\theta)$, which we observe commute with each other, since $e^{i\alpha}$ is a

scalar. Thus, we can consider these two components of U to be carried out in either order, or even simultaneously if we prefer.

Let's start by looking at $R_{\hat{n}}(\theta)$. At first, we might guess that the worst-case effort that is required to perform $R_{\hat{n}}(\theta)$ for angles θ where $-\pi \leq \theta \leq \pi$ ought to just turn out to be $|\theta|/2$, since, for example, a Bloch sphere rotation through an angle of $\theta = \pi$ radians corresponds to inverting a spin in ordinary 3D space through an angle of 180° to point in the opposite direction, which is an orthogonalizing transformation, and we already know from the Margolus–Levitin theorem that any transition to an orthogonal state under a constant Hamiltonian requires a minimum action (given zero ground state energy) for the state in question of $h/4 = (\pi/2)\hbar = (\pi/2)\,\text{rad}$, or an area swept out of $\pi/4$ square units. This is a good first guess, but later, we will see that the actual worst-case action turns out to be twice as large as this. (Our intuition forgot to take into account the fact that the state vector in the Margolus–Levitin theorem isn't actually the worst-case one, as far as the accumulated Hamiltonian action is concerned.)

Indeed, for any real unit 3-vector $\hat{n}$ (the "axis of rotation" for the Bloch sphere), one can easily verify that there is always a corresponding complex state vector

$$|v_{\hat{n}}^{+}\rangle = \frac{1}{\sqrt{2(1+n_z)}} \begin{bmatrix} n_z + 1 \\ n_x + i n_y \end{bmatrix} \tag{63}$$

which is a unit eigenvector of $\hat{n} \cdot \boldsymbol{\sigma}$ having eigenvalue +1. This state vector is therefore also an eigenstate of $R_{\hat{n}}(\theta)$, with eigenvalue $e^{i(\theta/2)}$. In other words, in any orthonormal basis that includes $|v_{\hat{n}}^{+}\rangle$ as one of the basis vectors, as θ increases from 0 (for now, we'll assume for simplicity that the final value of θ is non-negative, $0 \leq \theta \leq \pi$), the coefficient of the $|v_{\hat{n}}^{+}\rangle$ component of the state $|\psi(t)\rangle = R_{\hat{n}}(\theta)|v_{\hat{n}}^{+}\rangle$ (starting from the initial state $|\psi(0)\rangle = |v_{\hat{n}}^{+}\rangle$, where the coefficient $c_{|v_{\hat{n}}^{+}\rangle}$ is 1) describes a circular arc in the complex plane centered on the origin, sweeping out a total angle of $\theta/2$, and an origin-centered area of $\theta/4$. As we saw earlier, this same measure of the weighted-average accumulated angle and total area accumulated still holds in any basis. So, we have that the effort of $R_{\hat{n}}(\theta)$ must be at least $\theta/2$. Indeed, this is the exact worst-case effort, since $|v_{\hat{n}}^{+}\rangle$'s eigenvalue is maximal, so no pure energy eigenstate can possibly sweep out a larger angle as θ increases, and therefore no superposition of energy eigenstates (i.e., no general state) can do so either.

Now, what about the $e^{i\alpha}$ factor that's included in the expression for a general $U \in U_2$? Note that this term represents an overall (global) phase factor that applies to all eigenstates. As such, even the ground state $|g\rangle$

of whatever Hamiltonian is used to implement U might still accumulate a phase due to this phase factor. In this case, $|g\rangle$ would have non-zero Hamiltonian energy. If we redefine $|g\rangle$ to instead have zero energy ($H|g\rangle = 0$), then $|g\rangle$'s coefficient would not phase-rotate at all, since the action operator $A = Ht$ would give $A|g\rangle = 0$ for this state, and $U|g\rangle$ would give $(e^{iA})|g\rangle = (e^0)|g\rangle = |g\rangle$, that is, $|g\rangle$ would be unchanged by this U. However, it does not follow that we can always just let α be zero, as $|g\rangle$ may generally have accumulated an additional phase resulting from the $R_{\hat{n}}(\theta)$ component of U as well. It is the *total* phase accumulated by the ground state that we wish to define to be zero.

Let us now consider the following: Under the transformation $R_{\hat{n}}(\theta)$, as θ increases from 0, we notice that $|v_{\hat{n}}^+\rangle$ (the eigenvalue-1 eigenstate of $\hat{n}\cdot\boldsymbol{\sigma}$ which we constructed above) only phase-rotates by an angle $\theta/2$. Under $U = e^{i\alpha} R_{\hat{n}}(\theta)$, $|v_{\hat{n}}^+\rangle$ therefore undergoes an overall phase-rotation by an angle of $\alpha + \theta/2$. We confidently conjecture that the "least potential action" or most efficient way to implement U is to apply a Hamiltonian that simultaneously sweeps both α and θ forward steadily from 0, at respective rates that are exactly proportional to their intended final values. If this is correct, then $|v_{\hat{n}}^+\rangle$ is indeed an eigenstate of that best-case Hamiltonian, with energy $(\alpha + \theta/2)/t$ (recall that we're using $\hbar = 1$), where t is the total time taken for α and θ to reach their final values.

However, since the space we are working with is two-dimensional, there must be another energy eigenstate as well. Solving the eigen-equation $(\hat{n}\cdot\boldsymbol{\sigma})|v\rangle = r|v\rangle$, we find that the other eigenvalue r of $\hat{n}\cdot\boldsymbol{\sigma}$ is -1, and the other unit-length eigenvector, modulo phase-rotations, is (for $n_z > 0$)

$$|v_{\hat{n}}^-\rangle = \frac{1}{\sqrt{2(1-n_z)}} \begin{bmatrix} n_z - 1 \\ n_x + in_y \end{bmatrix} \tag{64}$$

or, in the special case when $n_z = 0$, then instead any normalized column vector $|v_{\hat{n}}^-\rangle = [v_0; v_1]$ where $|v_0| = |v_1| = 2^{-1/2}$ will work, so long as the vector components v_0 and v_1 have the specific obtuse (that is, $> 90°$) relative phase angle that is given by the relation $v_1 = (-n_x - in_y)v_0$. (Note that $|n_x + in_y| = 1$ when $n_z = 0$.)

Thus, for any Hamiltonian that smoothly sweeps θ forward in a steady transformation $R_{\hat{n}}(\theta)$ with $\theta \propto t$, there will actually be two different energy eigenstates having energies that are negatives of each other, one state in which the accumulated action of the Hamiltonian is $\theta/2$ (as we saw above), and another state (the ground state) where the action is the negative of this, or $-\theta/2$. Together with the global phase-rotation of α, we have that the total action for U is $\alpha + \theta/2$ and $\alpha - \theta/2$ for these two energy eigenstates, respectively.

Following our convention that the total action in the ground state should be always considered to be zero, we can shift the energy levels upwards in such a way that the lower value $\alpha - \theta/2$ will be equal to 0, in other words, we can adjust our rate of global phase rotation (which determined α) in such a way that we have exactly $\alpha = \theta/2$. Now, the total action in the high energy state is $\alpha + \theta/2 = \theta/2 + \theta/2 = \theta$.

In other words, starting with any $U \in \mathrm{U}_2$ and decomposing it as $U = \mathrm{e}^{\mathrm{i}\alpha} R_{\hat{n}}(\theta)$, which involves a rotation of the Bloch sphere through an angle of θ about an axis $\hat{n}$, we can calculate a meaningful difficulty $\mathcal{D}^+(U)$ by using the convention that the ground state should be considered to have energy 0, and by letting $\mathcal{D}^+(U) = \mathcal{D}^+(U_{\hat{n}}(\theta))$, where we define $U_{\hat{n}}(\theta) \equiv \mathrm{e}^{\mathrm{i}\theta/2} R_{\hat{n}}(\theta)$, that is, ignoring the original value of α (whatever it was) and instead adjusting α to have the value $\alpha = \theta/2$ which assigns the ground state to zero energy. Thus, we can say that the "true" computational/physical difficulty of U (given this choice) is exactly θ for any single-qubit unitary $U = \mathrm{e}^{\mathrm{i}\alpha} R_{\hat{n}}(\theta)$, regardless of the value of α. If θ is a pure number (implicitly bearing an angle unit of radians), then the worst-case Hamiltonian action to carry out the desired transform using the best-case Hamiltonian (assuming that is indeed what we have managed to characterize above) is $\theta\hbar$, in whatever physical units we wish to express $\hbar$. That is, $\mathcal{D}^+(U) = \theta$.

To wrap up this section, let us take a look at the precise form of the Hamiltonian that we are proposing. Note that

$$\hat{n} \cdot \boldsymbol{\sigma} = \begin{bmatrix} n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & -n_z \end{bmatrix} \tag{65}$$

is itself an Hermitian operator which plays the role of the Hamiltonian operator H with respect to the Bloch-sphere rotation unitary $R_{\hat{n}}(\theta) = \mathrm{e}^{\mathrm{i}(\theta/2)(\hat{n}\cdot\boldsymbol{\sigma})}$, if the rotation angle θ is taken be equal to twice the time t. Meanwhile, in this scenario, the extra phase-rotation factor $\mathrm{e}^{\mathrm{i}\alpha} = \mathrm{e}^{\mathrm{i}(\theta/2)}$ out front corresponds simply to an additional constant energy of +1, using the same angular velocity units of $(\theta/2t)$. This gives us a total "Hamiltonian" (in quotes because we haven't introduced an explicit time parameter here yet) of $H_{\hat{n}}$ that is required to implement a steady rotation about $\hat{n}$ which is equal to

$$\begin{aligned} H_{\hat{n}} &= 1 + \hat{n} \cdot \boldsymbol{\sigma} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & -n_z \end{bmatrix} \\ &= \begin{bmatrix} 1 + n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & 1 - n_z \end{bmatrix}. \end{aligned} \tag{66}$$

With this choice of "Hamiltonian," we can easily check that the $|v_{\hat{n}}^{\pm}\rangle$ are indeed its energy eigenstates, with $H_{\hat{n}}|v_{\hat{n}}^{-}\rangle = 0$ (the ground state has "energy" 0) and $H_{\hat{n}}|v_{\hat{n}}^{+}\rangle = 2$, which is what we want since it will cancel out with the 2 in the denominator of the exponent in the rotation unitary $U_{\hat{n}}(\theta) = e^{i\theta/2} R_{\hat{n}}(\theta) = e^{i(\theta/2)(1+\hat{n}\cdot\sigma)} = e^{i(\theta/2)H_{\hat{n}}}$.

To generalize the picture slightly, if a rotation through θ about an axis $\hat{n}$ is to take place over an arbitrary amount of time t, then we require a Hamiltonian (a proper one now, in actual angular-velocity energy units) of

$$H = \frac{\theta}{2t} H_{\hat{n}} = \frac{\theta}{2t} \begin{bmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{bmatrix} \tag{67}$$

With this choice of Hamiltonian, note that things work out nicely so that the high-energy eigenstate $|v_{\hat{n}}^{+}\rangle$ phase-rotates at exactly the desired rate $\omega^{+} = \theta/t$, since we have that

$$H|v_{\hat{n}}^{+}\rangle = \frac{\theta}{2t} H_{\hat{n}}|v_{\hat{n}}^{+}\rangle = \frac{\theta}{2t} 2|v_{\hat{n}}^{+}\rangle = \frac{\theta}{t}|v_{\hat{n}}^{+}\rangle = \omega^{+}|v_{\hat{n}}^{+}\rangle. \tag{68}$$

Thus, the action operator $A = Ht$ comes out exactly equal to the angle operator Ω which gives the total angle of phase rotation for both the energy eigenstates $|v_{\hat{n}}^{\pm}\rangle$, that is, $A|v_{\hat{n}}^{-}\rangle = \Omega|v_{\hat{n}}^{-}\rangle = 0|v_{\hat{n}}^{-}\rangle$ and $A|v_{\hat{n}}^{+}\rangle = \Omega|v_{\hat{n}}^{+}\rangle = \theta|v_{\hat{n}}^{+}\rangle$. And for an arbitrary initial state ψ, i.e., for any normalized complex superposition of the eigenstates $|v_{\hat{n}}^{\pm}\rangle$, $A[\psi] = \Omega[\psi]$ gives the quantum mean angle of phase rotation.

Note that in all the above discussion, we have assumed that the rotation angle is non-negative, i.e., that $0 \le \theta \le \pi$ (rad). To complete the picture, note that for values of θ between 0 and $-\pi$, we can convert them to positive angles by the simple expedient of rotating instead by an angle of $|\theta| = -\theta$ about the $-\hat{n}$ axis , which is an exactly equivalent rotation. This has the effect of exchanging the values of the $|v_{\hat{n}}^{\pm}\rangle$ eigenstates, as well as the sign of the $H_{\hat{n}}$ component of H. Other than that, everything else is the same, with the result that the action A always comes out non-negative and equal to the absolute value of θ. Of course, for the case of absolute angles outside the range $(-\pi, \pi]$, we can just reduce them to the equivalent angle in $(-\pi, \pi]$ by adding or subtracting the appropriate multiple of 2π.

In the above, although we have not yet quite finished proving rigorously that the specific H we have given is in fact the one that implements U with the least possible value of the worst-case action A, still, we expect that it should already seem highly plausible to the reader that this should in fact be the case, due to the directness and simplicity of our construction, which made use only of the simple fact that any arbitrary $U \in \mathrm{U}_2$ can be decomposed into a single generalized rotation about an arbitrary axis

is real three-space, accompanied by a global phase rotation. Of course, a more complete proof of the optimality of this construction would be desirable to have, but it will have to wait for future work.

12.2. Specific Single-Qubit Gates

Given the above discussion, to determine the difficulty $\mathcal{D}$ of any single-qubit gate U is a simple matter of finding some unit 3-vector $\hat{n}$ and angles $\alpha, \theta \in (-\pi, \pi]$ such that $U' = e^{i\alpha} R_{\hat{n}}(\theta)$, which is always possible. This then establishes that $\mathcal{D}^+(U) = |\theta|$, under our ground zero energy convention. Let us look briefly at how this calculation comes out for various single-qubit gates of interest.

1. The Pauli spin-operator "gates" $X = \sigma_x$ (which is the in-place NOT operation in the computational basis), $Y = \sigma_y$, and $Z = \sigma_z$ all of course involve a rotation angle of $\theta = \pi$, since they all square to the identity (2π rotation). Thus, $\mathcal{D}^+(X) = \mathcal{D}^+(Y) = \mathcal{D}^+(Z) = \pi = h/2$.

2. The "square root of NOT" gate $N = \frac{1}{2}\left[\begin{smallmatrix} 1+i & 1-i \\ 1-i & 1+i \end{smallmatrix}\right]$ of course requires an angle of $\pi/2$, since $N^2 = X$. Thus, $\mathcal{D}^+(N) = \pi/2 = h/4$.

3. The Hadamard gate $N = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ requires a rotation angle of π about the $\hat{n} = (1, 0, 1)/\sqrt{2}$ axis, i.e., $\hat{n} \cdot \boldsymbol{\sigma} = (\sigma_x + \sigma_z)/\sqrt{2}$. Also note that $H^2 = 1$ and a rotation through 2π is the identity. Thus, $\mathcal{D}^+(H) = \pi = h/2$.

4. The "phase gate" $S = \left[\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right]$ requires $\theta = \pi/2$ since note that $S^2 = Z$. So, $\mathcal{D}^+(S) = \pi/2 = h/4$.

5. The so-called "$\pi/8$" gate $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \exp[i\pi/4] \end{smallmatrix}\right]$ involves $\theta = \pi/4$ since note that $T^4 = Z$. Thus, $\mathcal{D}^+(T) = \pi/4 = h/8$.

6. The generalized phase gate $\mathrm{ph}(\theta) = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \exp[i\theta] \end{smallmatrix}\right]$ is just a rotation by an angle of θ about the z axis, so $\mathcal{D}^+(\mathrm{ph}(\theta)) = \theta = \theta\hbar$.

As a point of comparison, the paper [16] studies the time required to perform the specific gate $U = e^{i\theta} X$ (i.e., NOT with global phase rotation) using an optimal Hamiltonian, and conclude that the minimum time τ required (for a specific initial state) is

$$\tau = \frac{h}{4E}\left(1 + 2\frac{\theta}{\pi}\right). \tag{69}$$

Note that the corresponding Hamiltonian action α or effort $\mathcal{F}$ is

$$\alpha = \mathcal{F} = E\tau = \frac{h}{4} + 2\frac{h}{4}\frac{\theta}{\pi}$$

$$= \frac{\pi}{2}\hbar + \theta\hbar$$
$$= \frac{\pi}{2} + \theta \quad (\text{with } \hbar = 1). \tag{70}$$

At first glance, this might appear to contradict our claim that the difficulty of such a U ought to be exactly π. However, we should keep two things in mind. First, in Ref. 16, Levitin *et al.*, are concerned with the time to carry out U in the case of a specific subset of initial states which will actually transition to an orthogonal state in the time τ. However, these particular states are not the "worst-case" ones from our perspective, and so they don't determine the maximum effort. Rather, the particular states under consideration in their paper all have a mean energy of only $\bar{E} = (E_1 + E_2)/2$, where E_1 and E_2 are the low and high energy eigenvalues of the ideal Hamiltonian, respectively. Letting $E_1 = 0$ (our ground zero assumption), we have that $E_2 = 2\bar{E}$. Since E_2 has the highest energy available given this spectrum, the E_2 energy eigenstate accumulates more action over the time τ than any other possible state, in particular, double that of states with energy $\bar{E} = E_2/2$, and thus it is the E_2 state that determines the worst-case action, which is twice that of[16] or in other words $A = \pi$. The term involving θ in (70) drops out entirely, since as we already saw earlier, global phase shifts are irrelevant when considering total action, under our convention that the ground state action is always defined to be zero. Levitin *et al.* don't make this adjustment, because they are assuming that the Hamiltonian has already been arranged in advance to have a desired energy scale. Thus, the global phase rotation by θ leads to an extra additive θ in their expression (70) for the action.

12.3. Difficulty of Achieving Infidelity

A natural and widely-used measure of the degree of closeness or similarity between two quantum states u, v is the *fidelity*, which is defined (for pure states) as $F(u, v) = |\langle u|v\rangle| = |u^\dagger v|$. (see Ref. 23.) Note that if the actual state of a system is u, and we measure it in a measurement basis that includes v as a basis vector, the square of the fidelity $p = F^2$ gives the probability that the measurement operator will project the state down to v, and that v will be seen as the "actual" state. (This is a "quantum jump" or "wavefunction collapse" event, or, in the many-worlds picture, it is the subjectively experienced outcome when the state of the observer becomes inextricably entangled with that of the system.) Likewise with the roles of u and v reversed. Thus, only when $F = 0$ are the states u and v orthogonal.

We can also define a related quantity, the "infidelity" $\text{Inf}(u, v) \equiv \sqrt{1-p} = \sqrt{1-F^2}$. The squared infidelity between u and v is then just the

probability $1-p$ that if the actual state is u, then it will *not* be taken to v by a projective measurement (in a measurement basis that includes v), and vice-versa. In other words, if v is some old state of a system, and u is its new state, the squared infidelity between u and v is the probability that the answer to the question "Is the state different from v yet?" will be found to be "yes" when this question is asked experimentally by a measurement apparatus that compares the state with v.

Let us now explore the minimum effort that is required in order for some of the possible state vectors of a system to attain a given degree of infidelity (relative to their initial states), in the case of two-dimensional Hilbert spaces. Note that not all vectors will achieve infidelity; in particular, the eigenvectors of any time-independent Hamiltonian will always have 0 infidelity.

We start by recalling from earlier that any 2-dimensional unitary can be considered a rotation of the Bloch sphere about some axis in ordinary (real-valued) 3-D space. Since a simple change of basis suffices to transform any axis to any other, we can without loss of generality presume a rotation about the z axis, represented by

$$R_{\hat{z}}(\theta)=\begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}. \tag{71}$$

We saw earlier that the effort of any such rotation (under the ground-zero convention) is always exactly θ. What initial state will gain infidelity most rapidly under this transformation? Until we figure this out, let us allow the initial state to be a general unit vector $|v\rangle=[v_0; v_1]=v_0|0\rangle+v_1|1\rangle$ in the basis $|0\rangle, |1\rangle$. Then $|u\rangle=R_{\hat{z}}(\theta)|v\rangle=[e^{-i\theta/2}v_0; e^{i\theta/2}v_1]$ as a column vector of complex coefficients. Now the fidelity between v and u is

$$\begin{aligned} F(v,u) &= |\langle v|u\rangle| = \left|\langle v|R_{\hat{z}}(\theta)|v\rangle\right| \\ &= \left|v_0^* e^{-i\theta/2}v_0 + v_1^* e^{i\theta/2}v_1\right| \\ &= \left|e^{-i\theta/2}|v_0|^2 + e^{i\theta/2}|v_1|^2\right| \\ &= \left|\left[\cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\right]|v_0|^2 + \left[\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right]|v_1|^2\right| \\ &= \left|\left(\cos\frac{\theta}{2}\right)(|v_0|^2+|v_1|^2) + i\left(\sin\frac{\theta}{2}\right)(|v_1|^2-|v_0|^2)\right| \\ &= \left|\left(\cos\frac{\theta}{2}\right) + i\left(\sin\frac{\theta}{2}\right)(|v_1|^2-|v_0|^2)\right|, \end{aligned} \tag{72}$$

where in the last line we have made use of the fact that $|v_0|^2+|v_1|^2=1$ for a normalized v. Now, F^2 is the sum of the squared real and imaginary

components of the expression inside the outermost absolute-value delimiters || above:

$$\begin{aligned}[F(u,v)]^2 &= \Im^2[\langle v|u\rangle]+\Re^2[\langle v|u\rangle] \\ &= \cos^2\left(\frac{\theta}{2}\right)+\sin^2\left(\frac{\theta}{2}\right)\left(|v_1|^2-|v_0|^2\right)^2 \\ &= \cos^2\left(\frac{\theta}{2}\right)+\sin^2\left(\frac{\theta}{2}\right)\left(1-4|v_1|^2|v_0|^2\right) \\ &= 1-4\sin^2\left(\frac{\theta}{2}\right)|v_1|^2|v_0|^2, \end{aligned} \tag{73}$$

where in getting from the second to the third line, we have again made use of the fact that $|v_0|^2+|v_1|^2=1$. We can reassure ourselves that the last line of (73) is always in the range [0,1], since $|v_0|^2|v_1|^2 \leq 1/4$ given that $|v_0|^2+|v_1|^2=1$. Note also that the fidelity is minimized when $|v_0|^2=|v_1|^2=\frac{1}{2}$, that is, when the two z-basis states are in an equal superposition. This is then the "worst case" (worst in terms of "least fidelity") which we wish to focus on.

So now, the infidelity $I=\mathrm{Inf}(u,v)=\sqrt{1-F^2(u,v)}$ comes out to be a reasonably simple expression:

$$\begin{aligned}\mathrm{Inf}(u,v) &= \sqrt{1-[F(u,v)]^2} \\ &= \sqrt{4\sin^2\left(\frac{\theta}{2}\right)|v_1|^2|v_0|^2} \qquad (74) \\ &= 2\left(\sin\frac{\theta}{2}\right)|v_0||v_1|. \qquad (75)\end{aligned}$$

Note that for any given angle of rotation in $0<\theta<\pi/2$, the infidelity is maximized when $|v_0|=|v_1|=1/\sqrt{2}$. For such v, we have $|v_0||v_1|=\frac{1}{2}$ and so

$$\mathrm{Inf}(u,v)=\sin\frac{\theta}{2}. \tag{76}$$

Thus, if we wish that some system initially in state v should achieve a desired degree I of infidelity (relative to its initial state) using a transformation of minimum effort, we must choose a unitary transformation that is a rotation $R_{\hat{n}}(\theta)$ about an axis $\hat{n}$ that is "perpendicular" to v, and rotate by an angle $\theta=2\cdot\arcsin(I)$. The Hamiltonian action α accumulated by "worst-case" (that is, maximum-energy) vectors under this transformation is (by definition) the difficulty $\mathcal{D}^+(R_{\hat{n}}(\theta))$ of that unitary, and is given by $\alpha=2\cdot\arcsin(I)$.

However, the specific initial vector v that we are dealing with will not have the maximum energy E (relative to ground) but rather half of this, or $E/2$, since half of its probability mass will be in the high-energy state, and half in the zero-energy ground state. Therefore, v's total Hamiltonian action (amount of change) along its trajectory will instead be exactly $\alpha(v) = \arcsin(I)$, a wonderfully simple expression. This α is the effort exerted by the specific state v as it traverses a maximally efficient path for achieving infidelity $I = \sin\alpha$.

So, for example, suppose we want to cause some given initial state v to transition to a new state that has only a probability of at most $p = 1/2$ of being confused with the initial state if it were measured. This is to say that the infidelity between the states should be at least $I = \sqrt{1-p} = 1/\sqrt{2}$, which requires the state to traverse a trajectory that has a length of at least $\theta = \arcsin(I) = \arcsin(1/\sqrt{2}) = \pi/4 = h/8$, which can be done using a minimum-difficulty unitary transform whose worst-case effort is twice as great as this, or $\pi/2 = h/4$, meaning that the worst-case (maximum-energy) states of the system would traverse a trajectory of this (greater) length under an optimal implementation of such a transformation.

Assuming that the actual given initial state in question is assigned an average energy of only E above the ground state, it will take time at least $t = h/8E$ to carry out a unitary transformation on this state that achieves a probability above $1/2$ of distinguishing it from the resulting state; whereas, if we are given that the *maximum* energy state in the qubit spectrum has energy E, then it will take time at least $t = h/4E$ to carry out the transform.

In other words, to carry out an operation in time t that yields a 50% probability (or less) of conflation of some initial states with their successors requires that the initial states in question must have energy at least $E = h/8t$, and that states of energy at least $E = h/4t$ must exist in the spectrum.

Note that the above results are also perfectly consistent with the Margolus–Levitin theorem.[5] That is, plugging in an infidelity of $I = 1$ to represent a transition to an orthogonal state, we find that the specific initial state's effort $\mathcal{F}(v) = \arcsin(1) = \pi/2$ while the worst-case difficulty for this transform is $\theta = 2\arcsin(1) = \pi$; these figures are twice that for the previous example. And so for a state to attain a 0% probability of conflation (i.e., to reach an orthogonal state) requires that it have at least twice the energy as the previous scenario, or $E = \pi/2t = h/4t$ (under the Hamiltonian used to carry out the transformation), while other energy levels of at least $\pi/t = h/2t$ must be present in the spectrum of the Hamiltonian operator being used.

12.4. Higher-dimensional Operations

Naturally, we are interested not only in unitaries in U_2, but also in higher dimensions, in particular, unitaries in the groups U_{2^n}, which correspond to general "quantum logic gate" operations (really, arbitrary quantum computations) operating on sets of n qubits.

In particular, let us focus on the "controlled-U" gates with one target bit, which take the general form (modulo qubit reorderings)

$$U' = C^{n-1}U \equiv \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & U \end{bmatrix} \tag{77}$$

where we have $2^n - 2$ ones along the diagonal, and a rank-2 unitary matrix U in the lower-right corner. In other words, for computational basis states $|b_0 b_1 \ldots b_{n-1}\rangle$, whenever the first $n-1$ qubits $b_0 b_1 \ldots b_{n-2}$ are not all 1's, the state remains unchanged; otherwise, the unitary U is performed on the final qubit b_{n-1}.

We observe immediately that $\mathcal{D}^+(U') \geq \mathcal{D}^+(U)$, since all the input states that undergo any change at all will undergo the exact same transformation (in the subspace associated with the last qubit) that they would if U were just applied unconditionally. Thus, the worst-case trajectories when conditionally applying U can be no shorter than the worst-case unconditional trajectories (under an optimal implementation).

Furthermore, if U by itself would be optimally implemented by the Hamiltonian H, then it is easy to believe that U' would likewise be optimally implemented by the Hamiltonian

$$H' = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & H \end{bmatrix} \tag{78}$$

that is, with 0's everywhere except for a copy of H in the lower-right 2×2 submatrix. It is easy to verify that this H', when exponentiated, indeed produces the desired U'. And since its worst-case difficulty is equal to our lower bound $\mathcal{D}^+(U)$, it is in fact an optimal H', assuming our earlier conjecture about the optimality of H is correct. In this case, if H' is actually an available Hamiltonian in the context one is considering, then the effort of U' is indeed exactly the same as the effort of U.

We can see from this example that when we consider the full space of mathematically describable Hamiltonians, we are likely to greatly

underestimate the effort, compared to what can actually be implemented. The typical known implementations of U' in terms of small local quantum gates would require a number of orthogonalizing operations that is at least linear in n, whereas in our case above, the effort is constant (upper-bounded by π). It seems likely that the effort for a physically realistic (e.g., field-theory based) Hamiltonian for this class of Us would have to be more than constant, since the interaction of n qubits to determine an outcome would appear to necessarily be a non-local process.

In most physical situations of interest, we will not necessarily have available Hamiltonians that are of any form desired, such as the form H' suggested above. Instead, we may only have available a more limited, perhaps parameterized suite of Hamiltonians, perhaps ones that are formed by a sum or time-sequence of specific, controllable, localized couplings having (say) at most 2 qubits each, as is popularly represented in the quantum computing literature using the schematic notation of quantum logic networks.

Obviously, whenever our space of available Hamiltonians is more restricted than the simple "all Hermitian operations" scenario analyzed above, the resulting values of $\mathcal{D}^+(U)$ will in general become much larger, and probably also much more difficult for us to analytically calculate. To compute $\mathcal{D}^+(U)$ for Hamiltonians that can plausibly be constructed within the context of particular experimental frameworks that are readily physically realizable in the lab (or in a manufactured product, e.g., a someday-hopefully-to-be-realized commercial quantum computer) is clearly a much more complex and difficult task than we have attempted to tackle in this paper. To address this problem more fully will have to wait for future work.

Still, we hope that the present work can at least serve as a fruitful conceptual foundation on which we can proceed to build meaningful analytical and/or numerical analyses of the physical/computational "difficulty" of performing various quantum operations. We also hope that this work will serve as a helpful stepping stone for future investigators who wish to continue exploring the many deep and rich interconnections between physical and computational concepts.

12.5. Classical Reversible and Irreversible Boolean Operations

Although in the above discussion we have focused on the effort required to carry out quantum gate operations, it is easy to extend the results to classical logic operations as well. Any classical reversible operation is just a special case of a quantum gate where the matrix elements of the unitary operator (in the computational basis) are 0 or 1. For example,

a reversible Toffoli gate or Controlled-Controlled-NOT (CCNOT) is a special case of the C^2U gate addressed in §12.4 above. Specifically, since the U in question is X (NOT), which has a rotation angle of π, the effort required for Toffoli must be at least π, and indeed is exactly π if arbitrary Hamiltonians can be constructed. Toffoli is a universal gate for classical reversible computation, so a construction of any classical reversible circuit out of Toffoli gates sets an upper bound (as a multiple of π) on the difficulty of that computation, apart from any extra effort that may be required to control transitions between gates (which could be substantial, but is probably close to linear in the number of operations performed).

As for ordinary irreversible Boolean operations, these can be embedded into reversible operations as follows. Consider, for example, a standard boolean inverter, whose function is irreversible as it is normally specified in an electrical engineering context. The explicit function of an inverter is to destructively overwrite its output node with the logical complement of its input. (Please note that this function is distinct from that of a classical reversible NOT operation, which simply toggles a bit in-place.) Due to Landauer's principle, the physical information contained in the output node cannot actually be destroyed, but is instead transferred to reside in the environment. So, we can model the ordinary inverter's function as a sequence of reversible operations as follows:

1. Exchange output bit with an empty bit in the device's environment
2. Increment an "environment pointer" to refer to the next empty bit in some unbounded list
3. Perform a CNOT between input node and (now empty) output node

The first step can be understood as the emission from the device of the old stored value of the bit, in the form of entropy. The second step can be viewed as implementing the continuous flow of entropy away from the device, to make room for discarding the results of subsequent inverter operations. Finally, the third step carries out the desired logical function. The above breakdown is not necessarily the simplest possible implementation of the classical inverter (although it is probably close), but it at least sets an upper limit on the number of quantum operations that are absolutely required.

The first step can be carried out by a unitary SWAP operation between the two bits in question. The second step can be carried out by an annihilate/create pair of operations that moves a "particle" by one position to point to the next empty location in the environment; this corresponds to a unitary operation that increments the state vector $|i\rangle$ of some subsystem that specifies the integer location i of the environment

pointer. Finally, the third step is just an ordinary CNOT, with an effort of π. In principle, we could calculate and add up the effort for all these steps, together with the effort needed to update a part of the machine state that keeps track of which step we are on, to arrive at an upper bound on the effort required to implement a classical inverter operation. However, this calculation might not be very meaningful unless we did more work to specify a detailed physical setup that would allow us to confirm that such a bound was achievable in a practical hardware implementation.

13. RELATION TO BERRY PHASE

An interesting question to ask about our quantity $\mathcal{F}$ is what relationship (if any) it has to the classic notion of the geometric or Berry phase of a quantum trajectory.[24–31] So far, the relationships between these concepts are not completely clear, and working them out in more detail will have to wait for future work. However, some initial remarks are in order.

Let $H(t)$ be any time-dependent Hamiltonian that implements the unitary U for t going from 0 to τ, and let $|\psi\rangle$ be an eigenvector of U, with eigenvalue $e^{i\phi}$. The state $|\psi\rangle$ thus undergoes a cyclic evolution in the projective (phase-free) Hilbert space. Aharonov and Anandan[26] point out the relation $-\phi = \alpha - \beta$ (the integrated form of their Eq. (2)), where α is the integral of the instantaneous Hamiltonian energy of the state,

$$\alpha = \frac{1}{\hbar}\int_{t=0}^{\tau} \langle\psi(t)|H(t)|\psi(t)\rangle \mathrm{d}t \tag{79}$$

and β is a term given by

$$\beta = \int_{t=0}^{\tau} \langle\tilde{\psi}(t)|\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\tilde{\psi}(t)\rangle \mathrm{d}t, \tag{80}$$

where $\tilde{\psi}(t)$ is any continuously gauge-twiddled version of $\psi(t)$ such that $\tilde{\psi}(0) = \tilde{\psi}(\tau) = \psi(0)$. Aharonov and Anandan's paper[26] revolves around their claim that this β quantity is a generalized version of the Berry phase that applies even to non-adiabatic evolutions.

However, if the results of the present paper are correct, then Aharonov and Anandan's β is always an arbitrary value congruent to 0 (modulo 2π) and thus is not a physically meaningful quantity. The reason is that the α in (79) is exactly our $\alpha = A[\psi(0)]$, where $U = e^{-iA}$ (in the usual sign convention, which A&A are using), and thus $\psi(0)$ is also an eigenvector of A with eigenvalue α, so $|\psi(\tau)\rangle = U|\psi(0)\rangle = e^{-i\alpha}|\psi(0)\rangle$. Since we are already

given that $\psi(\tau) = e^{i\phi}\psi(0)$, it follows that $\phi \equiv -\alpha \pmod{2\pi}$; thus $\beta \equiv 0 \pmod{2\pi}$. Any desired multiple of 2π can always be selected for β by appropriate choice of the function $\tilde{\psi}(t)$. So, β does not contain any information at all about the specific evolution $\psi(t)$, and thus it is not a physically meaningful quantity.

It it interesting to note that the A&A paper[26] never actually shows that their quantity β can ever be different from 0 (mod 2π), although they do prove that β has some other "interesting" properties (such as being independent of the gauge of the original trajectory) which of course are true trivially if β is always congruent to zero.

Thus, it seems that one implication of our results (assuming they are correct) is that Aharonov and Anandan's particular version (at least) of the "geometric phase" is a chimera, and does not really exist. Further study is needed to verify this conclusion more rigorously, and also to determine whether other definitions of the Berry phase might escape from it, and retain a useful physical meaning that relates in some way to our quantity α. Since many researchers have reported the experimental detection of Berry-type phases (e.g., see Ref. 32), it seems highly unlikely that our results will turn out to nullify all versions of the geometric phase for all quantum evolutions. However, as of this writing, the correct resolution of the apparent discrepancy between theory and experiment on this question is not yet clear.

14. CONCLUSION

In this paper, we have shown that any continuous trajectory of a normalized state vector can be measured by a real-valued quantity which we call the *effort* $\mathcal{F}$, which is given by the line integral, along the trajectory, of the imaginary component of the inner product between adjacent states along the trajectory. This quantity is basis-independent, and is numerically equal to the probability-weighted average phase angle accumulated by the basis state coefficients (in radians), and to twice the area swept out by the coefficients in the complex plane, and also to the action of the time-dependent Hamiltonian along the trajectory, in units of $\hbar$. This notion of effort can be easily extended to apply also to transformation trajectories $U'(t)$ over time, as well as to an overall resulting unitary transform U, where it measures the difficulty $\mathcal{D}$ or minimum effort (over available trajectories) required to implement the desired transform in the worst case (maximizing over the possible initial states). Our framework can be used to easily rederive a variety of related results obtained by earlier papers for various more specialized cases.

The major implication of these results is that there is indeed a very definite sense in which we can say that the physical concept of energy does indeed precisely correspond to the computational concept of the rate of computation, that is, we can validly say that energy *is* the rate of physical computing activity, defined as the rate of change of the state vector, according to the measure that we have described in this paper. Furthermore, we can validly say that physical action *is* (an amount of) computation, defined as the total amount of change of the state vector, in the sense we have defined.

What about different specific types of energy, and specific types of action? Later papers along this line of research will survey how different types of energy and action can validly be identified with computational activity that is engaged in different types of processes. For example, heat may be identified with energy whose detailed configuration information is unknown (is entropy), rest mass–energy can be identified with energy that is engaged in updating a system's internal state in its rest frame, potential energy with phase rotation due to emission/absorption of virtual particles, and so forth. As a preview, it turns out that we can even make our computational interpretation consistent with special relativity by subdividing the energy of a moving body (in a given observer frame) into the *functional* energy Φ that is associated with updating the body's internal state (this turns out to be just the negative Lagrangian $-L = H - pv$) and a *motional* part $M = pv$ (related to but not quite the same as kinetic energy) that is associated with conveying the body through space; relativistic momentum then turns out to be the motional computational effort exerted per unit distance traversed. Future papers will elaborate on these related themes in more depth.

It is hoped that the long-term outcome of this line of thought will be to eventually show how *all* physical concepts and quantities can be rigorously understood in a well-defined mathematical framework that is also simultaneously well-suited for describing physical implementations of desired computational processes. That is, we seek an eventual unifying mathematical foundation that is appropriate for not only physical science, but also for device-level computer engineering and for physics-based computer science. We expect that such a unifying perspective should greatly facilitate the future design and development of maximally efficient computers constructed from nanoscale (and perhaps, someday, even smaller) components, machines that attempt to harness the underlying computational resources provided by physics in the most efficient possible fashion.

ACKNOWLEDGMENTS

Thanks are due to Charles Bennett, Oscar Boykin, Maureen Kelly, Lev Levitin, Tom Toffoli, Charles Weatherford, Colin Williams and the anonymous reviewers for helpful comments on early drafts and/or discussions about these ideas.

REFERENCES

1. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379, 623 (1948).
2. R. Landauer, *IBM J. Res. Dev.* **5**, 183 (1961).
3. C. H. Bennett, *IBM J. Res. Dev.* **17**(6), 525 (1973).
4. C. H. Bennett, *Sci. Am.* **295**(5), 108 (1987).
5. N. H. Margolus and L. B. Levitin, *Physica D* **120**, 188, (1998); quant-ph/9710043.
6. S. Lloyd, *Nature* **406**, 1047 (2000); quant-ph/9908043.
7. S. Lloyd, *Phys. Rev. Lett.*, **88**(23), Jun. 10 (2002); article 237901Z, quant-ph/0110141.
8. S. Lloyd and Y. J. Ng, *Sci. Am.* **291**(5), 52 (2004).
9. R. Clausius, *Poggendorff's Ann.* **125**, 353.
10. L. Boltzmann, *Sitzungsberichte der Akademie der Wissenschaften, Wien, II*, **66**, 275, (1872). English translation in S.G. Brush, *Kinetic Theory*, Vol. 2, *Irreversible Processes*, (Pergamon Press, Oxford, 1966), pp. 88–175.
11. C. Cercignani, *Ludwig Boltzmann: The Man Who Trusted Atoms* (Oxford University Press, 1998).
12. M. Tegmark, *Ann. Phys.* **270**, 1 (1998).
13. P. Benioff, *Found. Phys.* **32**(7), 989 (2002).
14. A. Tyagi, In *Workshop on Physics and Computation, 1992. PhysComp '92*, Dallas, Texas, 2–4 October (IEEE Computer Society Press, 1992), pp. 262–266.
15. T. Toffoli, In Anthony Hey (ed.) *Feynman and Computation: Exploring the Limits of Computers*, (Perseus, 1998), pp. 348–392.
16. L. B. Levitin, T. Toffoli, and Z. Walton, arXiv:quant-ph/0210076, Oct. (2002).
17. V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. A* **67**, 052109 (2003).
18. V. Giovannetti, S. Lloyd, and L. Maccone, *Proc. SPIE* **5111**, 1 (2003).
19. G. Sterman, *An Introduction to Quantum Field Theory* (Cambridge University Press, 1993).
20. T. Needham, *Visual Complex Analysis* (Oxford University Press, 1997).
21. W. K. Wootters, *Phys. Rev. D* **23**(2), 357 (1981).
22. P. A. M. Dirac, *General Theory of Relativity* (Wiley, 1975). Reprinted in 1996 by Princeton University Press.
23. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
24. M. V. Berry, *Proc. Roy. Soc. Lond. A* **392**, 45 (1984).
25. B. Simon, *Phys. Rev. Lett.* **51**(24), 2167 (1983).
26. Y. Aharonov and J. Anandan, *Phys. Rev. Lett.* **58**(16), 1593 (1987).
27. J. Anandan and Y. Aharonov, *Phys. Rev. D* **38**(6), 1863 (1988).
28. J. Samuel and R. Bhandari, *Phys. Rev. Lett.* **60**(23), 2339 (1988).
29. J. Anandan and Y. Aharonov, *Phys. Rev. Lett.* **65**(14), 1697 (1990).
30. S. K. Bose and B. Dutta-Roy, *Phys. Rev. A* **43**(7), 3217 (1991).
31. J.Y. Zeng and Y.A. Lei, *Phys. Rev. A* **51**(6), 4415, (1995).
32. G. Falci, R. Fazio, G. Massimo Palma, J. Siewert, and V. Vedral, *Nature* **407**, 355 (2000).

Quantum Information Processing, Vol. 4, No. 4, October 2005 (© 2005)
DOI: 10.1007/s11128-005-7852-4

From Dirac to Diffusion: Decoherence in Quantum Lattice Gases

Peter J. Love[1–3] and Bruce M. Boghosian[1]

Received February 7, 2005; accepted July 10, 2005

We describe a model for the interaction of the internal (spin) degree of freedom of a quantum lattice-gas particle with an environmental bath. We impose the constraints that the particle-bath interaction be fixed, while the state of the bath is random, and that the effect of the particle-bath interaction be parity invariant. The condition of parity invariance defines a subgroup of the unitary group of actions on the spin degree of freedom and the bath. We derive a general constraint on the Lie algebra of the unitary group which defines this subgroup, and hence guarantees parity invariance of the particle-bath interaction. We show that generalizing the quantum lattice gas in this way produces a model having both classical and quantum discrete random walks as different limits. We present preliminary simulation results illustrating the intermediate behavior in the presence of weak quantum noise.

KEY WORDS: Quantum lattice gas; decoherence; quantum random walk.

PACS: 03.67.Lx; 05.40.Ca.

1. INTRODUCTION

Lattice gases are arguably the simplest models for the simulation of classical physical systems. These models provide elementary microscopic dynamics whose hydrodynamic limits are, *inter alia*, the diffusion equation, Burgers' equation and the Navier Stokes equations.[1] They also provide a simple arena for the creation of new models of physical phenomena such as multicomponent flow and dynamical geometry.[2–5] Lattice gases possess deterministic, stochastic and quantum (unitary) formulations.

[1]Department of Mathematics, Tufts University, Medford, Massachusetts, 02155, USA.
[2]Current Address: D-Wave Systems Inc., Suite 100, 4401 Still Creek Drive, Burnaby, British Columbia, Canada V5C 6G9.
[3]To whom correspondence should be addressed. E-mail: peter@dwavesys.com

Simulation of quantum systems on quantum computers remains one of the very few applications for which exponential speedup over classical computation is provable. The quantum lattice gases defined by Meyer[6] and by Boghosian and Taylor[7] may be simulated on a quantum computer in exponentially fewer steps than are required on a classical computer and with an exponential reduction in hardware. These models, together with other approaches by Lloyd[8,9] and by Ortiz and Gubernatis,[10,11] have made concrete Feynman's observation that simulation of quantum systems is hard on classical computers, but easy on quantum computers.[12–14]

Meyer was led to the definition of the quantum lattice gas by consideration of quantum generalizations of cellular automata (CA). The simplest possible quantum CA model would be a map in which the updated state of a cell depended linearly on the state of its neighbours and the global evolution rule was unitary. Meyer showed that the only such maps are trivial, namely, the identity map and the left or right shift, possibly multiplied by a phase. This "No-Go" theorem proves that there are no nontrivial homogeneous scalar unitary cellular automata.[15]

Quantum cellular automata models may evade this No-Go theorem[15] by having cell values which are not scalar, or by having local update rules which are not linear (although the global evolution in such models remains quantum mechanical, and therefore linear). Non-scalar models sacrifice some simplicity, while the determination of unitarity for nonlinear models is problematic.[16–18] A third possibility exists: We may *partition* our cellular automata, dividing our evolution into two substeps, acting on two distinct neighborhoods. This was the approach taken by Meyer[19] and also by Watrous.[20] If one of the substeps of the evolution is interpreted as propagation of the cell values to neighboring sites we may identify the partitioned cellular automata with a lattice gas model.[19]

The dynamics of all lattice gases take place by propagation of particles to neighboring sites on the lattice, followed by a local collision operation. In a stochastic lattice gas which obeys semi-detailed balance, the collision step is a doubly-stochastic Markov matrix acting on the state of a single lattice site. For a lattice gas with only a single particle moving on the lattice, the stochastic model reduces to a classical random walk. In the collision step of a quantum lattice-gas model, the state at a site is acted on by a unitary scattering matrix. Just as in the classical case, the one-particle sector of the model possesses an interpretation as a discrete-time discrete-space quantum random walk.

Quantum lattice gases are explicitly formulated as discrete models for quantum physics, and so are subject to additional physical constraints. In particular, the unitary scattering matrix is constrained to be parity

invariant. The quantum lattice gas was shown to yield the continuum propagator for the Dirac equation in one-dimension,[19] and also possesses the Schroedinger equation as a nonrelativistic continuum limit.[21]

The quantum random walk has also been studied extensively from a point of view quite different from that of physical modeling. Quantum random walks may provide an alternative route for the development of new quantum algorithms. Such walks have the property that the variance of the walk grows linearly with time, in sharp contrast to classical walks where the variance grows as the square root of time. This echoes the computational advantage of Grover's search algorithm, and indeed unstructured search can be reformulated as a quantum random walk problem. Discrete-time quantum walks have yielded exponential performance improvements in the hitting time on the hypercube, and continuous-time quantum walks have yielded an exponential improvement in the solution of a graph traversal problem. For an overview of the subject of quantum random walks we refer the reader to the review of Kempe.[22]

Two things distinguish quantum lattice gases from discrete quantum walks. Firstly, the absence of scalar homogeneous models means that discrete quantum walks must also introduce an extra degree of freedom. This degree of freedom is interpreted as a "coin" state which determines the motion of the walker at the next time step. In the context of the quantum lattice gas this degree of freedom is interpreted as the spin (or, in one-dimension, helicity) degree of freedom of the particle. Secondly, the constraint of parity invariance imposed by Meyer in one-dimension, and discrete rotation invariance imposed by Boghosian and Taylor in d-dimensions is infrequently applied to discrete time quantum random walks. The coin state is commonly updated by the Hadamard operation, which is not invariant under parity inversion. While parity inversion symmetry is a fundamental requirement for models of physics, from the point of view of the computational properties of quantum walks parity invariance is not an obvious requirement.

Unitary actions and measurements are the elementary operations allowed on closed quantum systems. However, no quantum system (excepting, possibly, the entire universe) is truly closed. Open quantum systems may be treated as subsystems of some larger closed quantum system. The unitary evolution of the entire closed system induces a (generally non-unitary) evolution on the open quantum subsystem. Such operations, which include measurements and unitary actions in their span, are referred to as *quantum operations* in the context of quantum information theory.[23] The theory of open quantum systems has been developed from the point of view of fundamental physics in the work of Feynman and Vernon, Caldeira and Leggett and Prokof'ev and Stamp.[24–27] In this seminal

work the environment degrees of freedom are included explicitly, and renormalization group arguments are adduced to model the environment as belonging to one of two universality classes: either the environment degrees of freedom are localized (spin bath) or delocalized (oscillator bath). "Random level" models of the environment have also been studied.[28,29]

In keeping with the motivation of this issue we shall leave the fascinating computational properties of both discrete-time and continuous-time quantum random walks aside. We shall focus instead on the physical interpretation of the classical stochastic lattice gas as a microscopic model for diffusion and of the quantum lattice gas as a microscopic model for a single particle obeying the Dirac equation in the continuum limit. Our aim in the present paper will be the construction of a single model capable of capturing both types of behavior in different parameter regimes.

We shall consider decoherence arising from an interaction with an environment coupling only to the particles' internal degrees of freedom and not to the particles' position degrees of freedom. Such decoherence models are referred to as "coin" decoherence in the context of quantum random walks. In general, coin decoherence produces a transition from behavior characteristic of a quantum walk in which the standard deviation of the walk grows linearly in time to behavior characteristic of a classical walk in which the standard deviation varies as the square root of time.[30–35] Previous decoherence models for quantum walks considered a scattering step which is a unitary action with probability $1 - p$, and a unitary action followed by a projective measurement with probability p.[30–33] Models in which measurements on the coin yield less than total information have also been considered.[36]

We begin by considering a decoherence model for quantum lattice gases (or equivalently, discrete-time quantum random walks) with an explicitly physical motivation. We require that the system-environment interaction be fixed and that the interaction preserve the parity invariance of the original lattice gas. We first describe the unitary lattice-gas dynamics and the extension of such dynamics to the density matrix formalism, necessary for the introduction of quantum operations into such dynamics. We then introduce the framework of quantum operations, and discuss two types of decoherence models for quantum lattice gases. We derive the constraints on the unitary operator coupling the system and environment arising from the requirement of parity invariance. Numerical results are presented for one parameterization of the decoherence model, which qualitatively verify the transition from quantum to classical (i.e., diffusive) behavior in the model. We close the paper with discussion and some directions for future work.

2. DECOHERENCE

The state vector formulation of quantum mechanics is inadequate to describe situations in which we have imperfect knowledge of the quantum state. Such perfect knowledge is expressed by the system being in a *pure state*, that is, a vector of complex amplitudes whose moduli squared sum to one. Consider the imperfect preparation of a quantum state such that state $|\psi\rangle$ is prepared with probability $P(|\psi\rangle)$, where the (classical) probabilities $P(|\psi\rangle)$ sum to one. We cannot simply represent this by a real linear convex combination of states, as we already have a complex superposition over observable basis states. We must instead consider a linear real convex combination of the dyads $|\psi\rangle\langle\psi|$,

$$\rho = \sum P(\psi)\,|\psi\rangle\langle\psi|\,. \tag{1}$$

Such a real linear convex combination is called a *mixed state*. The density matrix is the appropriate tool for simultaneously describing two probabilistic aspects of the theory—one arising from classical uncertainty about the state of the system, and one arising from the fundamental uncertainty arising from quantum superposition.

The time evolution of the density matrix may be obtained by linear extension of the evolution of the pure states. The time evolution of a pure state is given by $|\psi'\rangle = U|\psi\rangle$ where U is the unitary evolution operator of the system, and so the time evolution of the density matrix is given by conjugation: $U:\rho' = U\rho U^{\dagger}$, where dagger indicates the Hermitian conjugate. The Hilbert space of a quantum system which can be divided into two subsystems A and B possesses a basis which can be tensor factored such that each basis vector $|m\rangle$ can be expressed as a tensor product $|m_a\rangle \otimes |m_b\rangle$, such that $|m_a\rangle \in H_a$ and $|m_b\rangle \in H_b$. The reduced density matrix of subsystem A is obtained by taking the partial trace of the full density matrix over subsystem B.

We may now define a quantum operation on the density matrix of system A. We take the tensor product of the density matrix of the system A with that of the environment B. A unitary operation U^{AB} acts on the resulting density matrix by conjugation. The environmental subsystem B is then traced over, resulting in a new density matrix for the system A. Such quantum operations are therefore maps from density matrices to density matrices. Such maps may be constructed without reference to an environment state by invoking the operator-sum representation. The theory of such maps may also be formulated axiomatically without reference to the constructive procedure adduced here.[23] In the present paper we utilize the unitary representation of quantum operations given above, while noting

that a formulation of the noise model we shall construct in terms of the operator-sum representation is possible, and in fact may be a more convenient representation of the model.

3. QUANTUM LATTICE-GAS MODEL

In the present paper we restrict attention to one-dimensional lattice gases with two directions per site. The (classical) particle states are specified as follows: Each site on the lattice has two lattice vectors connecting it to its left and right neighbors. There may be at most one particle per site per vector.[37] The dynamics of all lattice gases take place in two substeps. First, the particles propagate along their vectors to neighboring sites, retaining their velocity as they do so. Second, the particles at each site undergo a collision changing the occupations of the vectors at each site. The propagation step clearly conserves any quantity that is obtained by summing a function of particle mass and velocity over all particles, since those quantities are not changed as particles propagate. The collision step is required to conserve a subset of these quantities that are of physical interest, such as mass, momentum, etc. In the following we shall consider models which conserve particle number only; if we regard the particles as each having unit mass, this may be thought of as conservation of mass.

For a one-dimensional lattice with two vectors per site, the only deterministic rules which preserve particle number are trivial. The first nontrivial model occurs when one considers a stochastic collision in which a single particle at a site has probability $1-p$ to reverse direction. The stochastic lattice gas with a single particle may be identified with a random walk. Generalizing to multiple particles we find the evolution of the single particle distribution function for a classical stochastic lattice gas obeys the diffusion equation. The stochastic models include the (trivial) deterministic models as the special case $p=1$.

The single timestep evolution operator U of the quantum lattice gas without decoherence is the composition of advection and scattering steps:

$$\begin{aligned} \sum \psi_{x,\alpha}|x,\alpha\rangle &\xrightarrow{\text{advect}} \sum \psi_{x,\alpha}|x+\alpha,\alpha\rangle \\ &\xrightarrow{\text{scatter}} \sum \psi_{x,\alpha} S_{\alpha\alpha'}|x+\alpha,\alpha'\rangle, \end{aligned} \tag{2}$$

where the scattering matrix may be parameterized up to a global phase as:

$$S=\begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix} \tag{3}$$

and the Hilbert space is $2N$-dimensional, where N is the lattice size. This quantum lattice gas was shown to yield the continuum propagator for the Dirac equation for a particle with mass $\tan\theta$ in Ref. 19. The model may also be interpreted as a discrete-space discrete-time quantum random walk, subject to the constraint that the scattering rule be parity invariant.

We generalize the above model by first extending the dynamics to those of the density matrix for the lattice gas. This allows us to handle the mixed states which will arise when we introduce decoherence into the dynamics. We couple the particles' internal degrees of freedom to a bath of arbitrary size and act on the internal degrees of freedom and the bath with a unitary matrix which is the product:

$$U_c \cdot [S \otimes \mathbb{I}_{bath}], \tag{4}$$

where $\mathbb{I}$ is the identity operator on the bath. We then trace over the environment degrees of freedom at each timestep.

This decoherence model corresponds to a generalization of the quantum lattice gas to the case where the collision operator is neither a unitary operator, nor a classical Markov matrix as in the stochastic classical lattice gas, but a quantum operation. The set of quantum operations contains both unitary actions and Markov actions as special cases. The unitary operator is clearly included as a special case when one considers U_c which do not couple the system and the environment (i.e., which tensor factor into $U_c = U_{c(sys)} \otimes U_{c(bath)}$).

It is clear that the Markov operations of the classical stochastic lattice gas are included when one considers an initialization of the density matrix of the lattice gas in a completely classical state – that is, a density matrix at each site with only diagonal entries, p_l and p_r. The action of the unitary part of the collision operator on this matrix is:

$$\begin{aligned}
&\begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p_r & 0 \\ 0 & p_l \end{pmatrix} \begin{pmatrix} \cos\theta & -i\sin\theta \\ -i\sin\theta & \cos\theta \end{pmatrix} \\
&\quad = \begin{pmatrix} p_l\cos^2\theta + p_r\sin^2\theta & i\sin\theta\cos\theta(p_r - p_l) \\ i\sin\theta\cos\theta(p_l - p_r) & p_l\sin^2\theta + p_r\cos^2\theta \end{pmatrix}
\end{aligned} \tag{5}$$

This reproduces the action of a classical Markov matrix on the vector of probabilities (p_l, p_r), where the probability that the particle continues in its current state is $\cos^2\theta$ if the interaction with the environment induces the map:

$$\begin{aligned}
&\begin{pmatrix} p_l\cos^2\theta + p_r\sin^2\theta & i\sin\theta\cos\theta(p_r - p_l) \\ i\sin\theta\cos\theta(p_l - p_r) & p_l\sin^2\theta + p_r\cos^2\theta \end{pmatrix} \\
&\qquad \mapsto \begin{pmatrix} p_l\cos^2\theta + p_r\sin^2\theta & 0 \\ 0 & p_l\sin^2\theta + p_r\cos^2\theta \end{pmatrix}.
\end{aligned} \tag{6}$$

We can construct this mapping by coupling the system to a two-dimensional environment in the completely mixed state $1/2\mathbb{I}$ using the controlled-NOT gate where the system is the control qubit. More generally, the map given in Eq. (6) is the effect of a measurement on the system, and such actions are included in the set of quantum operations. The correspondence of unitary operations combined with measurements as equivalent to Markov operations is also discussed in the context of Type-II quantum computing in Ref. 38.

In order to complete the demonstration that classical stochastic lattice gases are included as a subset of models described by a density matrix whose collision process is a quantum operation, we must show that a classical (i.e., diagonal) density matrix evolves to another classical density matrix under propagation. The classical density matrix is non-zero only in entries $|x,\alpha\rangle\langle x,\alpha|$, which evolve under propagation to $|x+\alpha,\alpha\rangle\langle x+\alpha,\alpha|$. Diagonal density matrices evolve to diagonal density matrices under propagation and a subset of collision operations evolving diagonal density matrices to diagonal density matrices are equivalent to the Markov matices implementing collisions of a classical stochastic lattice gas.

Hence a single-particle quantum lattice-gas simulation in which the entire density matrix is stored and in which the collision rule is a quantum operation includes as special cases quantum lattice-gas evolution and the stochastic evolution of a classical lattice gas. Deterministic classical lattice gases are included as they are a special case of stochastic lattice gases. Such a model therefore provides a bridge between the quantum lattice gas which possesses the Dirac equation as a continuum limit and classical stochastic lattice gases which possess the diffusion equation as a continuum limit.

4. PARITY-PRESERVING NOISE

We now define the set of quantum operations giving our new scattering rule. Specification of a set of quantum operations defines a "quantum noise" model. Such a model has two ingredients: a model for the environment state and a model for the system-environment interaction. Before specifying the environment state we must specify how many dimensions the environment Hilbert space must have. Here we may invoke a theorem which states that for a d-dimensional system Hilbert space a d^2-dimensional environment is sufficient to produce every possible quantum operation on the system. This theorem states that for a quantum operation specified by a set of principal components in the operator-sum representation it is always possible to find a unitary operator coupling a

d^2-dimensional environment to a d-dimensional system which reproduces this quantum operation.[23]

This suggests the following noise model: We initialize the four-dimensional environment in a fiducial state ($|00\rangle$ for example). We then sample from a distribution over the unitary group $U(8)$ and apply our sampled operator to the eight-dimensional system-environment pair, and then we trace over the environment. Such a model samples from the entire set of quantum operations. The preparation of the environment in a fiducial state does not imply a loss of generality, as the sampled operation can be considered to be first an operator acting only on the environment preparing a random environment state, followed by an operation coupling the system and environment. The distribution over the set of quantum operations is induced in a nontrivial way by the distribution over $U(8)$. Such a model would resemble a random level, or random matrix model, of the environment.[28,29]

On the other hand, our aim is the construction of a noise model with physically inspired constraints on the system-environment interaction. For such physically motivated noise models we wish the system-environment coupling matrix to be a constant unitary operator, arising from a putative fixed system-environment interaction Hamiltonian (which we may or may not know). If we regard our quantum lattice gas as a discrete model for the Dirac equation, we note that the interaction of the helicity degree of freedom of a Dirac particle with an environment is indeed fixed by fundamental physics.

We must take care about the meaning of the theorem invoked above for noise models with a constant system-environment interaction. The theorem does not state that a fixed unitary operator coupling a d^2-dimensional environment to a d-dimensional system can reproduce every quantum operation on the system. In the sequel we construct our noise model for an environment of arbitrary dimension, although we revert to a four-dimensional environment for reasons of computational tractability for simulations.

Equation (3) gives an explicit parameterization of all parity-preserving two-dimensional unitary operators, up to a global phase. Such convenient parameterizations of quantum operations do not yet exist, and so we follow the explicit constructive procedure for such operations given above. We choose a model for the environment such that its state is a unimodular complex vector whose Cartesian components are independent random variables. The environment-system interaction is fixed, and we consider a wide class of such interactions, namely those which preserve the parity invariance of the original unitary quantum lattice gas. The unitary update S obeys parity invariance $ST = TS$ where T is the parity exchange

operator. The parity preserving couplings U_S have the property that they commute with the parity exchange operator acting on the system tensored with the identity operator acting on the environment.

$$[U_S, T \otimes \mathbb{I}] = 0 \tag{7}$$

Equation (7) expresses a discrete symmetry of a discrete-time evolution operator. In physics we are more usually concerned with continuous symmetries and continuous time evolution operators. The usual statement of invariance of an interaction Hamiltonian under a particular symmetry transformation is that the infinitesimal generators of the symmetry transformation commute with the Hamiltonian. In the language of Lie groups this means that the Hamiltonian lies in the commutator subalgebra of the generators of the symmetry transformation in the Lie algebra of $U(N)$, where N is the number of degrees of freedom of our system. We may also apply these ideas to a discrete symmetry of a discrete-time discrete-space model. Let t be the Lie algebra element corresponding to $T \otimes \mathbb{I}$. Let u be the Lie algebra element corresponding to U. A sufficient condition that Eq. (7) holds is:

$$[u, t] = 0. \tag{8}$$

The Lie algebra of $U(N)$ is the set of anti-hermitian matrices. We use the basis arising from the root system of the Lie algebra of $U(N)$:[39]

$$\begin{aligned}
D^p_{kl} &= i\delta_{pk}\delta_{pl} && 1 \le p \le N \\
S^{qp}_{kl} &= i(\delta_{kp}\delta_{ql} + \delta_{kq}\delta_{pl}) && 1 \le p \le N \quad q < p \\
A^{qp}_{kl} &= (\delta_{kp}\delta_{ql} - \delta_{kq}\delta_{pl}) && 1 \le p \le N \quad q < p
\end{aligned} \tag{9}$$

where δ_{xy} is the Kronecker delta, and there is no sum on repeated indices. We note that $A^{qq} = 0$ and $S^{qq} = 2D^q$. The convention for the antisymmetric matrices is chosen so that the labelling superscripts increase from left to right, and so that the negative entry is always in the upper triangular portion of the matrix.

The commutation relations of the Lie algebra basis follow from Eq. (4):

$$\begin{aligned}
[D^p, D^q]_{kl} &= 0 \quad 1 \le p \le N \\
[D^r, S^{qp}]_{kl} &= \delta_{pr}A^{qr} + \delta_{rq}A^{pr} \\
[D^r, A^{qp}]_{kl} &= \delta_{rp}S^{rq} - \delta_{rq}S^{rp} \\
[S^{rs}, A^{qp}]_{kl} &= \delta_{sp}S^{rq} - \delta_{sq}S^{rp} + \delta_{rp}S^{sq} - \delta_{rq}S^{sp} \\
[S^{rs}, S^{qp}]_{kl} &= \delta_{ps}A^{qr} + \delta_{qs}A^{pr} + \delta_{pr}A^{qs} + \delta_{qr}A^{ps} \\
[A^{rs}, A^{qp}]_{kl} &= \delta_{ps}A^{qr} + \delta_{qs}A^{pr} + \delta_{pr}A^{qs} + \delta_{qr}A^{ps}
\end{aligned} \tag{10}$$

The block diagonal form of $T \otimes \mathbb{I}$ makes it straightforward to diagonalize, and it is therefore straightforward to obtain the Lie algebra element t.

$$t = \ln T \otimes 1 = \frac{\pi}{2} \sum_{r=1}^{N} D^r - \frac{\pi}{2} \sum_{s=1}^{N/2} S^{(2s-1)2s} \tag{11}$$

A general element u of the Lie algebra may be written:

$$u = \sum_{r=1}^{N} \alpha_r D^r + \frac{1}{2} \sum_{p=1}^{N} \sum_{q=1}^{N} [\beta_{qp} S^{qp} + \gamma_{qp} A^{qp}] \tag{12}$$

where $\beta_{qp} = \beta_{pq}$ and $\gamma_{qp} = -\gamma_{pq}$, and $\beta_{qq} = \gamma_{qq} = 0$.

The commutator is then

$$[u, t] = \frac{\pi}{2} \left[\sum_{s=1}^{N} D^s - \sum_{s=1}^{N/2} S^{(2s-1)2s}, \sum_{r=1}^{N} \alpha_r D^r + \frac{1}{2} \sum_{p=1}^{N} \sum_{q=1}^{N} [\beta_{qp} S^{qp} + \gamma_{qp} A^{qp}] \right] \tag{13}$$

Applying the structure constants of the Lie algebra and using $\beta_{qp} = \beta_{pq}$ and $\gamma_{qp} = -\gamma_{pq}$ gives:

$$\begin{aligned} [u, t] = -\frac{\pi}{2} \sum_{s=1}^{N/2} \sum_{p=1}^{N} \Big[& \beta_{p(2s)} A^{p(2s-1)} + \beta_{p(2s-1)} A^{p(2s)} \\ & + \gamma_{p(2s)} S^{(2s-1)p} + \gamma_{p(2s-1)} S^{(2s)p} \Big] \end{aligned} \tag{14}$$

The constraint that this be zero imposes a set of constraints on the β coefficients and a set of constraints on the γ coefficients. Because the coefficients and the matrices A are real, while the matrices S are pure imaginary, we may rearrange terms involving the A's and S's separately to obtain these constraints. We write

$$[u, t] = C_\beta + C_\gamma. \tag{15}$$

where $C_\beta = 0$ and $C_\gamma = 0$ are necessary conditions for $[u, t] = 0$.

$$C_\beta = -\frac{\pi}{2} \sum_{s=1}^{N/2} \sum_{p=1}^{N} \left[\beta_{p(2s)} A^{p(2s-1)} + \beta_{p(2s-1)} A^{p(2s)} \right]$$

$$= -\frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ even}^{N} \beta_{ps} A^{p(s-1)} - \frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ odd}^{N-1} \beta_{ps} A^{p(s-1)}$$

$$-\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ even}^{N} \beta_{ps} A^{p(s+1)} - \frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ odd}^{N-1} \beta_{ps} A^{p(s+1)} \tag{16}$$

We wish to rearrange terms so that we have a unique set of A's whose coefficients we can set to zero in order to obtain our constraints. In the first and fourth terms here the indices on the A's have opposite parity, whereas in the second and third terms the indices have the same parity in each term, but the indices are both odd in the second term and both even in the third term. This means that the first and fourth terms may be combined by exchanging dummy indices and using the symmetry properties of the A's, whereas the second and third terms must be dealt with separately. Denoting term x in the right hand side of Eq. (16), C_β^x, and taking the first and fourth terms:

$$C_\beta^1 + C_\beta^4 = -\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ even}^{N} \beta_{p(s+1)} A^{ps} - \frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ odd}^{N-1} \beta_{p(s-1)} A^{ps} \tag{17}$$

Exchanging p and s in the second term on the right hand side:

$$C_\beta^1 + C_\beta^4 = -\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ even}^{N} \left[\beta_{p(s+1)} A^{ps} + \beta_{s(p-1)} A^{sp} \right] \tag{18}$$

Using the antisymmetry of the A's, we have

$$C_\beta^1 + C_\beta^4 = -\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ even}^{N} \left[\beta_{p(s+1)} - \beta_{s(p-1)} \right] A^{ps} \tag{19}$$

Now consider the second and third terms in C_β. The second term is:

$$\begin{aligned} C_\beta^2 &= -\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ odd}^{N-1} \beta_{p(s+1)} A^{ps} \\ &= -\frac{\pi}{4} \sum_{s \ odd}^{N-1} \sum_{p \ odd}^{N-1} \left[\beta_{p(s+1)} - \beta_{s(p+1)}\right] A^{ps} \end{aligned} \tag{20}$$

The third term is:

$$\begin{aligned} C_\beta^3 &= -\frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ even}^{N} \beta_{p(s-1)} A^{ps} \\ &= -\frac{\pi}{4} \sum_{s \ even}^{N} \sum_{p \ even}^{N} \left[\beta_{p(s-1)} - \beta_{s(p-1)}\right] A^{ps} \end{aligned} \tag{21}$$

These equations give the following set of conditions on the beta coefficients:

$$\begin{aligned} &\beta_{p(s+1)} - \beta_{s(p-1)} = 0 \qquad s \mod 2 = 1 \quad p \mod 2 = 0 \\ &\beta_{p(s+1)} - \beta_{s(p+1)} = 0 \qquad s \mod 2 = 1 \quad p \mod 2 = 1 \\ &\beta_{p(s-1)} - \beta_{s(p-1)} = 0 \qquad s \mod 2 = 0 \quad p \mod 2 = 0 \end{aligned} \tag{22}$$

These constraints are redundant. In fact a pair of constraints is sufficient to ensure $C_\beta = 0$:

$$\begin{aligned} &\beta_{p(s+1)} - \beta_{s(p-1)} = 0 \qquad s \mod 2 = 1 \quad s < p \\ &\beta_{p(s+1)} - \beta_{s(p+1)} = 0 \qquad s \mod 2 = 1 \quad p \mod 2 = 1 \quad s < p \end{aligned} \tag{23}$$

We now consider C_γ.

$$\begin{aligned} C_\gamma &= -\frac{\pi}{2} \sum_{s=1}^{N/2} \sum_{p=1}^{N} \left[\gamma_{p(2s)} S^{(2s-1)p} + \gamma_{p(2s-1)} S^{(2s)p}\right] \\ &= -\frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ even}^{N} \gamma_{p(s-1)} S^{sp} - \frac{\pi}{2} \sum_{s \ even}^{N} \sum_{p \ odd}^{N-1} \gamma_{p(s-1)} S^{sp} \\ &\quad -\frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ even}^{N} \gamma_{p(s+1)} S^{sp} - \frac{\pi}{2} \sum_{s \ odd}^{N-1} \sum_{p \ odd}^{N-1} \gamma_{p(s+1)} S^{sp} \end{aligned} \tag{24}$$

Exchanging dummy indices in the third term, expanding the first and fourth terms, and utilizing the symmetry of the S's and the antisymmetry of the γ's gives:

$$\begin{aligned} C_\gamma = & -\frac{\pi}{4} \sum_{s\ even}^{N} \sum_{p\ even}^{N} \left[\gamma_{p(s-1)} + \gamma_{s(p-1)}\right] S^{sp} \\ & -\frac{\pi}{2} \sum_{s\ even}^{N} \sum_{s\ odd}^{N-1} \left[\gamma_{p(s-1)} - \gamma_{(p+1)s}\right] S^{sp} \\ & -\frac{\pi}{4} \sum_{s\ odd}^{N-1} \sum_{p\ odd}^{N-1} \left[\gamma_{p(s+1)} + \gamma_{s(p+1)}\right] S^{sp} \end{aligned} \tag{25}$$

This yields the constraints on the γ's:

$$\begin{aligned} &\gamma_{p(s-1)} - \gamma_{(p+1)s} = 0 \qquad s \mod 2 = 0 \quad p \mod 2 = 1 \quad p < s-1 \\ &\gamma_{p(s-1)} + \gamma_{s(p-1)} = 0 \qquad s \mod 2 = 0 \quad p \mod 2 = 0 \quad p \leq s \end{aligned} \tag{26}$$

5. SIMULATIONS

Simulations were performed for a system whose internal degree of freedom is coupled to a four-dimensional bath. The coupling operator is therefore an element of $U(8)$. The general form of an element of the Lie algebra of $U(8)$ obeying the constraints derived above is

$$i\begin{pmatrix} \alpha_1 & \beta_{12} & \beta_{13} & \beta_{14} & \beta_{15} & \beta_{16} & \beta_{17} & \beta_{18} \\ \beta_{12} & \alpha_2 & \beta_{14} & \beta_{13} & \beta_{16} & \beta_{15} & \beta_{18} & \beta_{17} \\ \beta_{13} & \beta_{14} & \alpha_3 & \beta_{34} & \beta_{35} & \beta_{36} & \beta_{37} & \beta_{38} \\ \beta_{14} & \beta_{13} & \beta_{34} & \alpha_4 & \beta_{36} & \beta_{35} & \beta_{38} & \beta_{37} \\ \beta_{15} & \beta_{16} & \beta_{35} & \beta_{36} & \alpha_5 & \beta_{56} & \beta_{57} & \beta_{58} \\ \beta_{16} & \beta_{15} & \beta_{36} & \beta_{35} & \beta_{56} & \alpha_6 & \beta_{67} & \beta_{57} \\ \beta_{17} & \beta_{18} & \beta_{37} & \beta_{38} & \beta_{57} & \beta_{67} & \alpha_7 & \beta_{78} \\ \beta_{18} & \beta_{17} & \beta_{38} & \beta_{37} & \beta_{58} & \beta_{57} & \beta_{78} & \alpha_8 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\gamma_{13} & -\gamma_{14} & -\gamma_{15} & -\gamma_{16} & -\gamma_{17} & -\gamma_{18} \\ 0 & 0 & -\gamma_{14} & -\gamma_{13} & \gamma_{16} & -\gamma_{15} & -\gamma_{18} & -\gamma_{17} \\ \gamma_{13} & \gamma_{14} & 0 & 0 & -\gamma_{35} & -\gamma_{36} & -\gamma_{37} & -\gamma_{38} \\ \gamma_{14} & \gamma_{13} & 0 & 0 & -\gamma_{36} & -\gamma_{35} & -\gamma_{38} & -\gamma_{37} \\ \gamma_{15} & \gamma_{16} & \gamma_{35} & \gamma_{36} & 0 & 0 & -\gamma_{57} & -\gamma_{58} \\ \gamma_{16} & \gamma_{15} & \gamma_{36} & \gamma_{35} & 0 & 0 & -\gamma_{58} & -\gamma_{57} \\ \gamma_{17} & \gamma_{18} & \gamma_{37} & \gamma_{38} & \gamma_{57} & \gamma_{58} & 0 & 0 \\ \gamma_{18} & \gamma_{17} & \gamma_{38} & \gamma_{37} & \gamma_{58} & \gamma_{57} & 0 & 0 \end{pmatrix} \tag{27}$$

In all cases the system was initialized in a pure state corresponding to a gaussian spatial wavefunction centred at the origin with a standard deviation equal to one quarter of the lattice size, with equal amplitudes for both internal states of the particle. The system was a periodic lattice with 64 sites, and the parameter θ in the unitary part of the collision operator was set equal to 0.35.

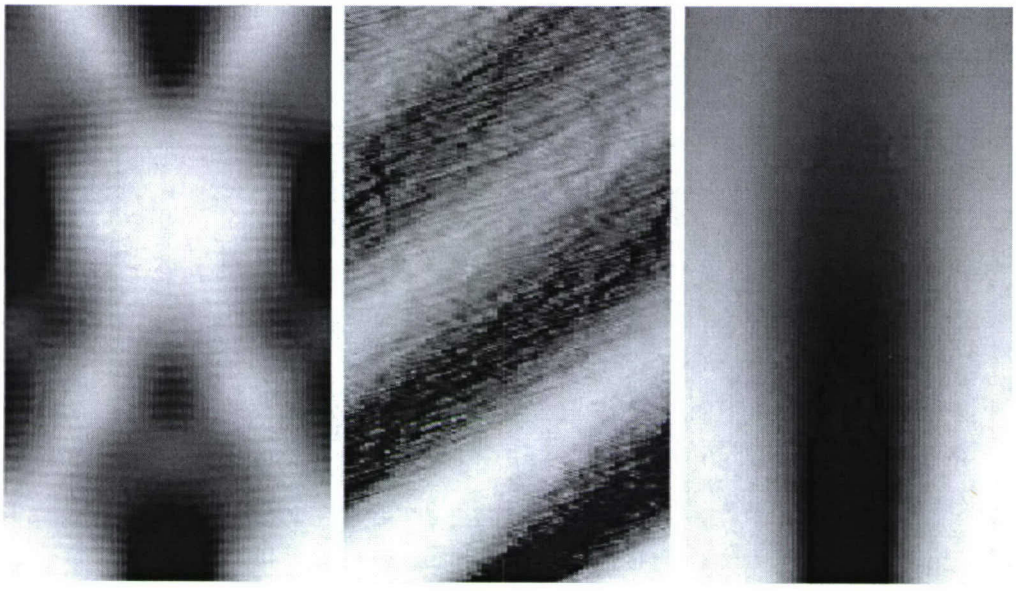


Fig. 1. Plots of the time evolution of the diagonal components for 400 time steps. The greyscale indicates the magnitude of the density matrix evolving from a Gaussian pure state centered at the origin with standard deviation one quarter of the system size. Left: Unitary evolution. Middle: Evolution with $\mathbb{Z}_8$ coupling the system to the bath. Right: Evolution with a parity-invariant coupling between system and bath.

Three types of simulations were performed. First, a simulation of the density matrix of the system and bath was performed in which the coupling matrix was the identity operator. In this case the quantum lattice gas reproduces the unitary evolution expected in the case that the system-bath coupling is zero. Second, simulations were performed in which the system-bath coupling operator is the Fourier transform over the cyclic group $\mathbb{Z}_8$. Finally, simulations were performed in which all coefficients α, β and γ in a matrix of the form (27) were set equal to one and the resulting matrix numerically exponentiated to obtain a parity-preserving coupling (Fig. 1).

The simulation with no system-bath coupling shows typical unitary evolution on a cycle, with the wave-packet dispersing until the edges of the wave-packet reach the periodic boundaries of the system, at which time interference occurs between the original packet and the reentrant components. The reversibility and unitarity of the dynamics is apparent as the system never settles into a static equilibrium state. The simulations in which the system-bath coupling is given by the Fourier transform, which violates parity invariance, exhibit a driving of the system to the right. Additional simulations in which the coupling is given by the conjugate of the Fourier transform with the parity inversion operator show the same driving effect in the opposite direction, as expected. Simulations in which

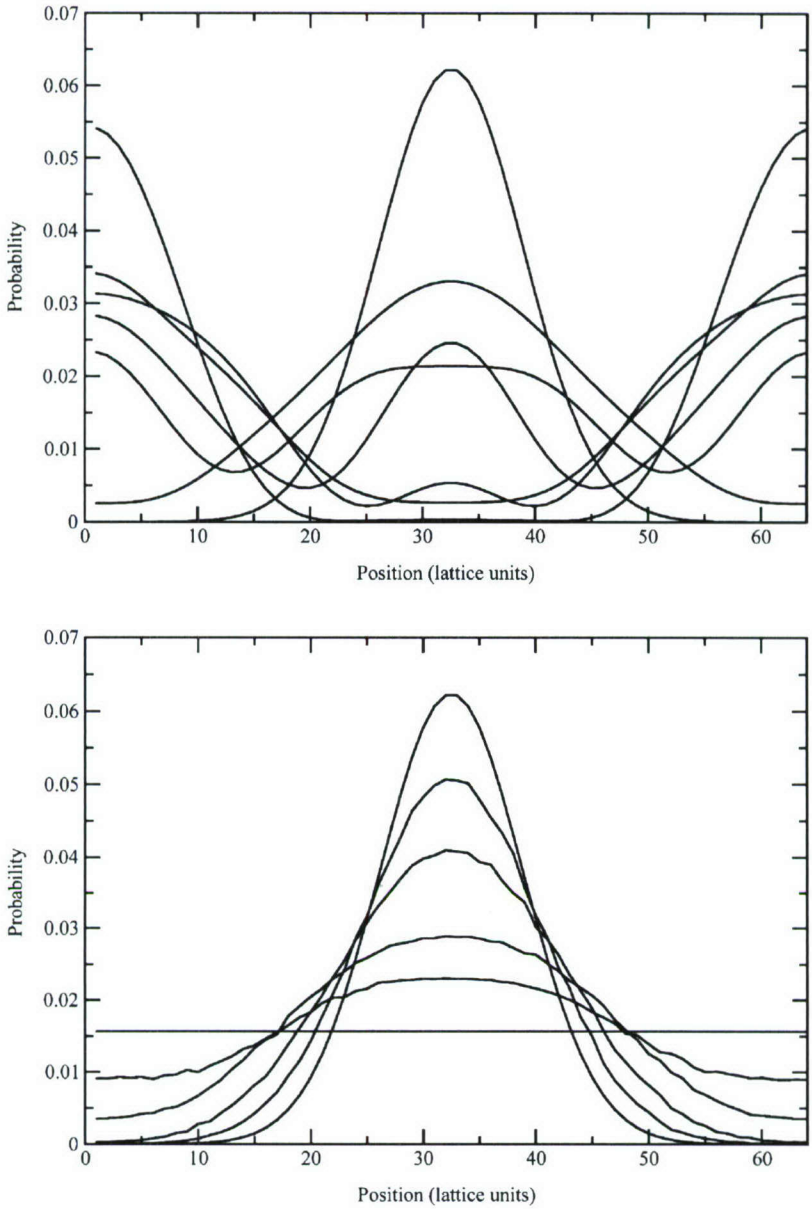


Fig. 2. Top: Time evolution of the probability distribution for the first 300 timesteps of unitary evolution. Curves are plotted every 50 timesteps. Bottom: Time evolution of the first 300 timesteps of the parity-invariant decoherent evolution and the final uniform distribution. Curves are plotted every 50 timesteps, and the final state after 8000 timesteps of evolution is shown.

the system is coupled to the bath by our example parity-invariant unitary matrix show no driving effect. The system undergoes an irreversible evolution of the initial probability distribution to the uniform distribution on the cycle.

The behavior of the model shows its intermediate quantum-classical nature in two ways. First, the decay of the probability distribution shows some residual "wave-like" behavior in addition to the overall damping. Second, the final density matrix is not a completely mixed state, but retains non-zero off-diagonal components within a finite band. Both these effects indicate that the coupling matrix chosen implements a coupling strong enough to cause irreversible dynamics on an observable timescale but weak enough that the dynamics retains some interesting quantum characteristics (Fig. 2).

6. CONCLUSIONS

We have defined a new quantum lattice-gas model for a single particle (or equivalently, a discrete-space discrete-time quantum random walk) in which the scattering rule is given by a quantum operation, rather than a unitary, deterministic, or stochastic operation. We showed that the model so defined includes unitary, stochastic and deterministic models as special cases, as well as interesting intermediate behavior. Preliminary simulation results confirm this by exhibiting non-unitary diffusive decay of an initial gaussian pure state.

The noise model chosen here utilized a fixed unitary operator coupling a four-dimensional bath to the two-dimensional Hilbert space of the internal degree of freedom of a single quantum lattice-gas particle. This distinguishes our work from the noise model given by Kendon and Sanders[36] in which a single environment qubit is coupled by an interaction with a strength tunable from the case of no coupling to the case where the environment produced a projective measurement of the coin degree of freedom of the quantum random walk.

As noted above, a two-dimensional environment with a fixed interaction is insufficient to reproduce all quantum operations. However, the constraints of parity invariance on the coupling operator were obtained here for $U(N)$, and so this work could be extended to include an arbitrarily large environment. If the most general noise model is desired, the sampling procedure discussed but not implemented above, in which randomly sampled unitary operators couple a four-dimensional bath to the internal degree of freedom of the particle, provably includes all quantum operations. The parameterization of the noise model then involves a parameterization of the distribution of quantum operations induced by a given distribution on the unitary group.

The above discussion motivates several directions for future work. First, the noise model presented here is not conveniently parameterized.

Ideally we would be able to smoothly vary the degree of coupling between the system and the environment from zero (where the time evolution would approximate the Dirac equation) to the case where the environment performs projective measurement of the particles' internal degrees of freedom. This property is possessed by the noise model of Kendon and Sanders,[36] and such a parameterization of the model presented here is certainly possible, although it may be tedious in practice. The work of Kendon and Sanders[36] discusses decoherence in quantum random walks from the point of view of complementarity. The work of the present paper was motivated instead by the principle of correspondence. The generalization here satisfies the requirement that the results agree with classical theory in the case that the particle is strongly coupled to an environment.

The two most natural generalizations of the work described here are to quantum lattice-gas models with multiple particles, and to models defined in multiple dimensions. The constraint of parity invariance becomes the constraint of invariance under discrete rotations in multiple dimensions, and one expects the constraints analogous to those derived here to be correspondingly more complex. Simulations of multiple particles, even in one dimension, have a classical computational cost which grows exponentially with the number of particles. However, one expects that few-particle simulations will be tractable.

The possibility of efficient quantum simulation of classical systems is an important open problem in the field of quantum computation for physical modeling. One of the central problems in this area is that most of the equations of classical physics of practical interest are irreversible macroscopic equations of motion. The work presented here shows that irreversibility may be simply included by the use of quantum operations instead of unitary matrices. The emergence of irreversible behavior in the degrees of freedom of the subsystem exhibited here is a manifestation of the "arrow of time" of non-equilibrium thermodynamics. The interesting practical question, which remains open, is whether there are systems for which the time complexity of such classical and quantum simulation is different. More glibly: Can time's arrow may be made to move faster on a quantum computer?

ACKNOWLEDGMENTS

PJL and BMB were supported by DARPA QuIST program administered under AFOSR grant number F49620-01-1-0566, and by ARO contract number W911NF-04-1-0334. BMB was also supported by AFOSR award number FA9550-04-1-0176. Both authors would like to thank

AFOSR for their hospitality at the Quantum Computation for Physical Modeling Workshop on Marthas Vineyard in 2004. The authors have great pleasure in thanking David Meyer, Gianluca Caterina, Howard Brandt, and Seth Lloyd for helpful discussions and questions.

REFERENCES

1. G. D. Doolen, U. Frisch, B. Hasslacher, S. Orszag, and S. Wolfram (eds.), *Lattice Gas Methods for Partial Differential Equations* (Addison-Wesley, 1990).
2. B. M. Boghosian, P. V. Coveney, and P. J. Love, *Proc. Roy. Soc. Lond. A* **456**, 1431 (2000); cond-mat/9907298.
3. P. J. Love, *Phil. Trans. R. Soc. Lond. Ser. A* **360**, 345 (2002); cond-mat/0109475.
4. P. J. Love, B. M. Boghosian, and D. A. Meyer, *Phil. Trans. R. Soc. Lond. Ser. A* **362**, 1667 (2004); cond-mat/0506742.
5. B. Hasslacher and D. A. Meyer, *Int. J. Mod Phys. C* **9**, 1597 (1998).
6. D. A. Meyer, *Phil. Trans. Roy. Soc. Lond. A* **360**, 395 (2002); quant-ph/0111069.
7. B. M. Boghosian and W. Taylor, Physica D **120**, 30 (1998).
8. S. Lloyd, *Science* **261**, 5128 (1996).
9. D. S. Abrams and S. Lloyd, *Phys. Rev. Lett.* **79**, 2586 (1997).
10. G. Ortiz, E. Knill, and J. E. Gubernatis, *Nucl. Phys. B (Proc. Suppl.)* **106**, 151 (2002).
11. G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, *Phys. Rev. A* **64**, 022319 (2001).
12. R. P. Feynman, *J. Opt. Soc. Am. B* **3**, 464 (1984).
13. R. P. Feynman, *Found. Phys.* **16**, 507 (1986).
14. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982).
15. D. A. Meyer, *Phys. Lett. A* **223**, 337 (1996).
16. C. Durr, Random Struct. Algor. **11**, 381 (1997).
17. C. Durr, *SIAM J. Comput.* **31**, 1076 (2002).
18. D. A. Meyer, quant-ph/9605023 (1996).
19. D. A. Meyer, *J. Stat. Phys.* **85**, 551 (1996).
20. J. Watrous, in *Proceedings of the 36th Annual Meeting on the Foundations of Computer Science* (1995), pp. 528–537.
21. B. M. Boghosian and W. Taylor, *Phys. Rev. E* **57**, 54 (1998).
22. J. Kempe, *Contemp. Phys.* **44**, 307 (2003); quant-ph/0303081.
23. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
24. R. P. Feynman and F. L. Vernon, *Ann. Phys.* **24**, 118 (1963).
25. A. O. Caldeira and A. Leggett, *Ann. Phys.* **149**, 374 (1983).
26. N. V. Prokof'ev and P. C. E. Stamp, *Rep. Prog. Phys.* **63**, 669 (2000).
27. U.Weiss, *Quantum Dissipative Systems* (World Scientific, 1993).
28. E. P. Wigner, *Ann. Math.* **53**, 36 (1953).
29. D. L. Hill and J. A. Wheeler, *Phys. Rev.* **89**, 1102 (1953).
30. G. Alagic and A. Russell, quant-ph/0501169 (2005).
31. V. M. Kendon and B. Tregenna, *Phys. Rev. A* **67**, 042315 (2002); quant-ph/0209005.
32. V. M. Kendon and B. Tregenna, quant-ph/0210047 (2002).
33. V. M. Kendon and B.Tregenna, in *Decoherence and Entropy in Complex Systems* (2003); Vol. 633, pp. 253–267, quant-ph/0301182.
34. T. A. Brun, H. A. Carteret, and A. Ambainis, quant-ph/0208195 (2002).

35. T. A. Brun, H. A. Carteret, and A. Ambainis, *Phys. Rev. A* **67**, 032304 (2002); quant-ph/0210180.
36. V. M. Kendon and B. C. Sanders, *Phys. Rev. A* **71**, 022307 (2004); quant-ph/0404043.
37. This is sometimes called the "exclusion principle" for lattice gases, but it should be noted that it is unrelated to the Pauli exclusion principle.
38. P. J. Love and B. M. Boghosian, *To appear in Physica A* (2005); quant-ph/0506244.
39. T. Brocker and T. tomDieck, *Representations of Compact Lie Groups* (Springer-Verlag, 1985).

Quantum Information Processing, Vol. 4, No. 5, November 2005 (© 2005)
DOI: 10.1007/s11128-005-0003-0

Unambiguous State Discrimination in Quantum Key Distribution

Howard E. Brandt[1]

Received March 8, 2005; accepted August 15, 2005

The quantum circuit and design are presented for an optimized entangling probe attacking the BB84 Protocol of quantum key distribution (QKD) and yielding maximum information to the probe. Probe photon polarization states become optimally entangled with the signal states on their way between the legitimate transmitter and receiver. Although standard von-Neumann projective measurements of the probe yield maximum information on the pre-privacy amplified key, if instead the probe measurements are performed with a certain positive operator valued measure (POVM), then the measurement results are unambiguous, at least some of the time. It follows that the BB84 (Bennett–Brassard 1984) protocol of quantum key distribution has a vulnerability similar to the well-known vulnerability of the B92 (Bennett 1992) protocol.

KEY WORDS: Quantum cryptography; quantum key distribution; quantum communication; quantum computer; entanglement.

PACS: 03.67.Dd, 03.67.Hk, 03.65.Ta.

1. INTRODUCTION

It is generally known that the standard two-state B92 protocol of quantum key distribution[1] is not a good solution for quantum key distribution.[2] Although the two nonorthogonal photon-polarization states of the protocol cannot in general be distinguished unambiguously by an eavesdropping probe without being disturbed, they can be unambiguously discriminated at least some of the time, at the cost of there occurring some inconclusive events. However, if the inconclusive rate equals the loss rate of the legitimate receiver (due to attenuation in the key distribution channel), and only the unambiguous states are relayed by the probe to the legitimate receiver, then the probe can obtain complete information on the

[1]U.S. Army Research Laboratory, Adelphi, MD, USA. E-mail: hbrandt@arl.army.mil

pre-privacy amplified key, once the bases are announced on the public channel during reconciliation. A POVM receiver is an implementation of such a probe.[3,4] Also to counter alteration in the attenuation due to the probe, the legitimate channel might be replaced by a more transparent one. In the present work, I argue that the four-state BB84 protocol of quantum key distribution[5] has a similar vulnerability.

In Section 2, the optimized unitary transformation is reviewed, representing the action of an optimized entangling probe on quantum key distribution in the BB84 protocol. In Section 3, the corresponding quantum circuit is obtained. In Section 4, a method is determined for projectively measuring the appropriate correlated states of the probe, thereby obtaining the maximum possible information (in the ideal case, ignoring losses). In Section 5, a possible implementation of the entangling probe is obtained. In Section 6, analysis is presented for alternatively using the Positive Operator Valued Measure (POVM) receiver in measuring the probe, thereby unambiguously discriminating the signal states, at least some of the time. Section 7 contains a summary.

2. OPTIMUM ENTANGLING PROBE

For the standard four-state BB84 protocol of quantum key distribution, the author recently calculated a simple optimized unitary transformation representing the effect of an eavesdropping probe, which on average yields the maximum information to the probe for a given error rate induced by the probe.[6] The most general possible probe consistent with unitarity was addressed,[7–12] in which each individual transmitted bit is made to interact with the probe so that the carrier and the probe are left in an optimum entangled state, and measurement of the probe then yields maximum information about the carrier state. The probe optimization was based on maximizing the information gain by the probe on corrected data for a set error rate induced by the probe in the legitimate receiver.[9–12] Corrected data include data remaining after discarding inconclusive results and also erroneous data as determined by block checksums and bisective search.[7,9,10] The maximum information gain by the probe is needed for privacy amplification.[7,10]

It was shown that the above unitary transformation representing the probe produces the following entanglements for initial probe state $|w\rangle$ and incoming BB84 signal photon-polarization states $|u\rangle$, $|\bar{u}\rangle$, $|v\rangle$, or $|\bar{v}\rangle$, respectively:[6]

$$|u\rangle \otimes |w\rangle \longrightarrow \frac{1}{4}\left(|\alpha_+\rangle \otimes |u\rangle + |\alpha\rangle \otimes |\bar{u}\rangle\right), \tag{1}$$

$$|\bar{u}\rangle \otimes |w\rangle \longrightarrow \frac{1}{4}\left(|\alpha\rangle \otimes |u\rangle + |\alpha_-\rangle \otimes |\bar{u}\rangle\right), \tag{2}$$

$$|v\rangle \otimes |w\rangle \longrightarrow \frac{1}{4}\left(|\alpha_-\rangle \otimes |v\rangle - |\alpha\rangle \otimes |\bar{v}\rangle\right), \tag{3}$$

$$|\bar{v}\rangle \otimes |w\rangle \longrightarrow \frac{1}{4}\left(-|\alpha\rangle \otimes |v\rangle + |\alpha_+\rangle \otimes |\bar{v}\rangle\right). \tag{4}$$

Here, the states $|u\rangle$ and $|\bar{u}\rangle$ are orthogonal signal states in the $\{|u\rangle, |\bar{u}\rangle\}$ basis, and $|v\rangle$ and $|\bar{v}\rangle$ are orthogonal signal states in the $\{|v\rangle, |\bar{v}\rangle\}$ basis, and the two bases are nonorthogonal with $\pi/4$ angle between the linear polarizations of states $|u\rangle$ and $|v\rangle$. The probe states $|\alpha_+\rangle$, $|\alpha_-\rangle$, and $|\alpha\rangle$ are given by

$$\begin{aligned}|\alpha_+\rangle = &\left[\left(2^{1/2}+1\right)(1+\eta)^{1/2} + \left(2^{1/2}-1\right)(1-\eta)^{1/2}\right]|w_0\rangle \\ &+\left[\left(2^{1/2}+1\right)(1-\eta)^{1/2} + \left(2^{1/2}-1\right)(1+\eta)^{1/2}\right]|w_3\rangle,\end{aligned} \tag{5}$$

$$\begin{aligned}|\alpha_-\rangle = &\left[\left(2^{1/2}-1\right)(1+\eta)^{1/2} + \left(2^{1/2}+1\right)(1-\eta)^{1/2}\right]|w_0\rangle \\ &+\left[\left(2^{1/2}-1\right)(1-\eta)^{1/2} + \left(2^{1/2}+1\right)(1+\eta)^{1/2}\right]|w_3\rangle,\end{aligned} \tag{6}$$

$$|\alpha\rangle = \left[-(1+\eta)^{1/2} + (1-\eta)^{1/2}\right]|w_0\rangle + \left[-(1-\eta)^{1/2} + (1+\eta)^{1/2}\right]|w_3\rangle, \tag{7}$$

where

$$\eta \equiv [8E(1-2E)]^{1/2}, \tag{8}$$

expressed in terms of the probe basis states $|w_0\rangle$ and $|w_3\rangle$, and also the set error rate E induced by the probe. Of particular importance is the fact that the states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ are in general nonorthogonal. The Hilbert space of the probe is two-dimensional, depending on the two probe basis vectors, $|w_0\rangle$ and $|w_3\rangle$. It should be noted here that in Eqs. (29) and (32) of Ref. 6, the overall signs must be positive in order to yield Eq. (19) of Ref. 6. Also, in Eqs. (5)–(7) above, the sign choices correspond to the implementation chosen here. Also, it is important here that large error rates are to be ignored to enforce monotonicity. The present analysis and that of Ref. 14, Eq. (71) of Ref. 13, and Eq. (89) of Ref. 6 apply for $E \leq 1/4$. [To include larger E, the optimization of Eq. (158) of Ref. 13 with $\sin 2\mu = 1$, $\cos\theta = 1$, and $p = (1-2E)^{1/2}$ can be implemented, also

with a single CNOT gate yielding the same maximum information gain and POVM measurement characteristics given here. However such large error rates are of less practical importance.]

It is to be noted in Eq. (1) that the projected probe state $|\psi_{uu}\rangle$ correlated with the correct received signal state (in the notation of Refs. 7, 12), in which the state $|u\rangle$ is sent by the transmitter, and is also received by the legitimate receiver, is $|\alpha_+\rangle$. Analogously, using Eq. (2), it follows that the correlated probe state $|\psi_{\bar{u}\bar{u}}\rangle$ is $|\alpha_-\rangle$. The two states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ are to be distinguished by the measurement of the probe. Also, according to Eqs. (3) and (4), the same two probe states are the appropriate correlated states $|\psi_{\bar{v}\bar{v}}\rangle$ and $|\psi_{vv}\rangle$, respectively. It will be shown below that the fact that the same two nonorthogonal states are to be distinguished in either polarization basis is critical to unambiguous state discrimination of the BB84 signal states (at least some of the time), since it is this feature of the entangling probe that will enable measurement of the probe by the POVM receiver.[3]

It should also be stressed that the state $|\alpha\rangle$ in Eqs. (1)–(4) is only correlated with the cases where one of the four signal states is sent and the orthogonal state is received by the legitimate receiver, and this will be interpreted as error. However, again it is to be remembered that the erroneous bits are discarded, and so their only role in the entangling probe optimization is in determining the error rate induced by the probe, and the latter is appropriately set by the eavesdropper. For any eavesdropping protocol, there is always a tradeoff between information gain by a probe and the induced error rate, and any key distribution protocol has some intrinsic noise level that the eavesdropper can exploit. It must also be understood that privacy amplification is always necessary, and this process is very error-rate dependent. Analogous arguments apply to attacks on the B92 protocol, which as stated above has a well-known POVM attack vulnerability.

3. QUANTUM CIRCUIT

In this Section, I exploit the quantum circuit model of quantum computation to determine the quantum circuit corresponding to the optimum unitary transformation, Eqs. (1)–(4). It was shown in Ref. 6 that the tensor products of the initial state $|w\rangle$ of the probe with the orthonormal basis states, $|e_0\rangle$ and $|e_1\rangle$, of the signal transform as follows (See Eqs. (1), (35)–(40) of Ref. 6), again with the above sign choices:

$$|e_0 \otimes w\rangle \longrightarrow |e_0\rangle \otimes |A_1\rangle \tag{9}$$

and

$$|e_1 \otimes w\rangle \longrightarrow |e_1\rangle \otimes |A_2\rangle\,, \tag{10}$$

expressed in terms of probe states $|A_1\rangle$ and $|A_2\rangle$, where

$$|A_1\rangle = a_1\,|w_0\rangle + a_2\,|w_3\rangle\,, \tag{11}$$

$$|A_2\rangle = a_2\,|w_0\rangle + a_1\,|w_3\rangle\,, \tag{12}$$

in which

$$a_1 = 2^{-1/2}(1+\eta)^{1/2}, \tag{13}$$

$$a_2 = 2^{-1/2}(1-\eta)^{1/2}, \tag{14}$$

and η is given by Eq. (8).

In the two-dimensional Hilbert space of the signal, the two orthogonal basis states $|e_0\rangle$ and $|e_1\rangle$ are oriented symmetrically about the signal states $|u\rangle$ and $|v\rangle$, and make angles of $\pi/8$ with the signal states $|u\rangle$ and $|v\rangle$, respectively. Next, consider a quantum controlled-NOT gate (CNOT gate), in which the control qubit consists of the two signal basis states $\{|e_0\rangle\,, |e_1\rangle\}$, and the target qubit consists of the probe basis states $\{|w_0\rangle\,, |w_3\rangle\}$, and such that when $|e_0\rangle$ enters the control port then $\{|w_0\rangle\,, |w_3\rangle\}$ becomes $\{|w_3\rangle\,, |w_0\rangle\}$ at the target output port, or when $|e_1\rangle$ enters the control port then $\{|w_0\rangle\,, |w_3\rangle\}$ remains unchanged. It then follows that a simple quantum circuit affecting the transformations (9) and (10), and thereby faithfully representing the entangling probe, consists of this CNOT gate with the state $|A_2\rangle$ always entering the target port, and $\{|e_0\rangle\,, |e_1\rangle\}$ entering the control port. When $|e_0\rangle$ enters the control port, then $|A_2\rangle$ becomes $|A_1\rangle$, or when $|e_1\rangle$ enters the control port then $|A_2\rangle$ remains unchanged, in agreement with Eqs. (9) and (10) with $|w\rangle = |A_2\rangle$. According to the quantum circuit model of quantum computation, it is known that at most three CNOT gates, together with single-qubit gates, are in general necessary and sufficient in order to implement an arbitrary number of unitary transformations of two qubits. Apparently in the present case, a single CNOT gate suffices to faithfully represent the optimized unitary transformation.

Next expanding the signal state $|u\rangle$ in terms of the signal basis states,[7] it then follows from Eqs. (9) and (10) that the CNOT gate affects the following transformation when the signal state $|u\rangle$ enters the control port:

$$|u\rangle \otimes |A_2\rangle \longrightarrow \cos\frac{\pi}{8}\,|e_0\rangle \otimes |A_1\rangle + \sin\frac{\pi}{8}\,|e_1\rangle \otimes |A_2\rangle\,. \tag{15}$$

Equivalently, then

$$|u\rangle \otimes |A_2\rangle \longrightarrow \cos\frac{\pi}{8}\left(\cos\frac{\pi}{8}|u\rangle - \sin\frac{\pi}{8}|\bar{u}\rangle\right)\otimes|A_1\rangle + \sin\frac{\pi}{8}\left(\sin\frac{\pi}{8}|u\rangle + \cos\frac{\pi}{8}|\bar{u}\rangle\right)\otimes|A_2\rangle, \tag{16}$$

or

$$|u\rangle \otimes |A_2\rangle \longrightarrow \frac{1}{4}[(2+2^{1/2})|A_1\rangle\otimes|u\rangle + (2-2^{1/2})|A_2\rangle\otimes|u\rangle - 2^{1/2}|A_1\rangle\otimes|\bar{u}\rangle + 2^{1/2}|A_2\rangle\otimes|\bar{u}\rangle], \tag{17}$$

which becomes Eq. (1). Analogously, it follows that the quantum circuit also yields Eqs. (2)–(4).

4. PROBE MEASUREMENT CORRELATIONS

According to Eqs. (1)–(4), and the analysis of Section 3, the probe produces the entanglements of Eqs. (1)–(4) for initial probe state $|w\rangle = |A_2\rangle$ and incoming BB84 signal states $|u\rangle$, $|\bar{u}\rangle$, $|v\rangle$, or $|\bar{v}\rangle$, respectively. Then if, following the public reconciliation phase of the BB84 protocol, the signal basis mutually selected by the legitimate transmitter and receiver is publicly revealed to be $\{|u\rangle, |\bar{u}\rangle\}$, then the probe measurement must distinguish the projected probe state $|\alpha_+\rangle$, when the signal state $|u\rangle$ is both sent and received, from the projected probe state $|\alpha_-\rangle$, when the signal state $|\bar{u}\rangle$ is both sent and received. In this case one has the direct correlations:

$$|u\rangle \Longleftrightarrow |\alpha_+\rangle, \quad |\bar{u}\rangle \Longleftrightarrow |\alpha_-\rangle. \tag{18}$$

The same two states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ must be distinguished, no matter which basis is chosen during reconciliation. Thus, according to Eqs. (3) and (4), if, following the public reconciliation phase of the BB84 protocol, the signal basis mutually selected by the legitimate transmitter and receiver is publicly revealed to be $\{|v\rangle, |\bar{v}\rangle\}$, then the probe measurement must distinguish the projected probe state $|\alpha_-\rangle$, when the signal state $|v\rangle$ is both sent and received, from the projected probe state $|\alpha_+\rangle$, when the signal state $|\bar{v}\rangle$ is both sent and received. In this case one has the direct correlations:

$$|v\rangle \Longleftrightarrow |\alpha_-\rangle, \quad |\bar{v}\rangle \Longleftrightarrow |\alpha_+\rangle. \tag{19}$$

Next, one notes that the correlations of the projected probe states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ with the probe's two orthogonal basis states $|w_0\rangle$ and $|w_3\rangle$

are indicated, according to Eqs. (5) and (6), by the following probabilities:

$$\frac{|\langle w_0|\alpha_+\rangle|^2}{|\alpha_+|^2} = \frac{|\langle w_3|\alpha_-\rangle|^2}{|\alpha_-|^2} = \frac{1}{2} + \frac{[E(1-2E)]^{1/2}}{(1-E)}, \tag{20}$$

$$\frac{|\langle w_0|\alpha_-\rangle|^2}{|\alpha_-|^2} = \frac{|\langle w_3|\alpha_+\rangle|^2}{|\alpha_+|^2} = \frac{1}{2} - \frac{[E(1-2E)]^{1/2}}{(1-E)}, \tag{21}$$

and one then has the following state correlations:

$$|\alpha_+\rangle \Longleftrightarrow |w_0\rangle\,, |\alpha_-\rangle \Longleftrightarrow |w_3\rangle\,. \tag{22}$$

Next combining the correlations (18), (19), and (22), one then establishes the following correlations:

$$\{|u\rangle\,, |\bar{v}\rangle\} \Longleftrightarrow |\alpha_+\rangle \Longleftrightarrow |w_0\rangle\,, \tag{23}$$

$$\{|\bar{u}\rangle, |v\rangle\} \Longleftrightarrow |\alpha_-\rangle \Longleftrightarrow |w_3\rangle\,, \tag{24}$$

to be implemented by the probe measurement method. This can be simply accomplished by a von Neumann projective measurement of the orthogonal probe basis states $|w_0\rangle$ and $|w_3\rangle$, implementing the probe projective measurement operators $\{|w_0\rangle\langle w_0|, |w_3\rangle\langle w_3|\}$. The chosen geometry in the two-dimensional Hilbert space of the probe is such that the orthogonal basis states $|w_0\rangle$ and $|w_3\rangle$ make equal angles with the states $|\alpha_+\rangle$ and $|\alpha_-\rangle$, respectively. This geometry is consistent with the symmetric von Neumann test, which is an important part of the optimization in Refs. 7, 9–12.

5. ENTANGLING PROBE DESIGN

Based on the above results, I have invented a simple entangling probe design.[15] An incident photon coming from the legitimate transmitter is received by the probe in one of the four signal-photon linear-polarization states $|u\rangle$, $|\bar{u}\rangle$, $|v\rangle$, or $|\bar{v}\rangle$ in the BB84 protocol. The signal photon enters the control port of the CNOT gate. The initial state of the probe is a photon in linear-polarization state $|A_2\rangle$ and entering the target port of the CNOT gate. The probe photon is produced by a single-photon source and is appropriately timed with reception of the signal photon by first sampling a few successive signal pulses to determine the repetition rate of the

transmitter. The photon linear-polarization state $|A_2\rangle$, according to Eqs. (12)–(14) and (8), is given by

$$|A_2\rangle = \left[\frac{1}{2}\{1-[8E(1-2E)]^{1/2}\}\right]^{1/2}|w_0\rangle + \left[\frac{1}{2}\{1+[8E(1-2E)]^{1/2}\}\right]^{1/2}|w_3\rangle, \tag{25}$$

and can be simply set for an error rate E by means of a polarizer. In this way the device can be tuned to the chosen error rate to be induced by the probe. The outgoing gated signal photon is relayed on to the legitimate receiver, and the gated probe photon enters a Wollaston prism, oriented to separate photon orthogonal-linear-polarization states $|w_0\rangle$ and $|w_3\rangle$, and the photon is then detected by one of two photodetectors. This is an ordinary von Neumann projective measurement. If the basis, revealed during the public basis-reconciliation phase of the BB84 protocol, is $\{|u\rangle, |\bar{u}\rangle\}$, then the photodetector located to receive the polarization state $|w_0\rangle$ or $|w_3\rangle$, respectively, will indicate, in accord with the correlations (23) and (24), that a state $|u\rangle$ or $|\bar{u}\rangle$, respectively, was most likely measured by the legitimate receiver. Alternatively, if the announced basis is $\{|v\rangle, |\bar{v}\rangle\}$, then the photodetector located to receive the polarization state $|w_3\rangle$ or $|w_0\rangle$, respectively, will indicate, in accord with the correlations (23) and (24), that a state $|v\rangle$ or $|\bar{v}\rangle$, respectively, was most likely measured by the legitimate receiver. By comparing the record of probe photodetector triggering with the sequence of bases revealed during reconciliation, then the likely sequence of ones and zeroes constituting the key, prior to privacy amplification, can be assigned. In any case the net effect is to yield, for a set error rate E, the maximum information gain to the probe, which is given by Eq. (19) of Ref. 6, namely,

$$I_{\text{opt}}^{R} = \log_2\left[2-\left(\frac{1-3E}{1-E}\right)^2\right]. \tag{26}$$

The geometry of the initial and shifted probe polarization states $|A_2\rangle$ and $|A_1\rangle$, respectively, and the orthogonal probe basis states, $|w_0\rangle$ and $|w_3\rangle$, in the two-dimensional Hilbert space of the probe, is such that the angle δ_0 between the probe state $|A_1\rangle$ and the probe basis state $|w_0\rangle$ is given by

$$\delta_0 = \cos^{-1}\left(\frac{\langle w_0|A_1\rangle}{|A_1|}\right), \tag{27}$$

or, substituting $|A_1\rangle$, given by Eqs. (11), (13), and (14), namely,

$$|A_1\rangle=\left[\frac{1}{2}\{1+[8E(1-2E)]^{1/2}\}\right]^{1/2}|w_0\rangle+\left[\frac{1}{2}\{1-[8E(1-2E)]^{1/2}\}\right]^{1/2}|w_3\rangle, \tag{28}$$

in Eq. (27), one obtains

$$\delta_0=\cos^{-1}\left(\frac{1}{2}\left\{1+[8E(1-2E)]^{1/2}\right\}\right)^{1/2}. \tag{29}$$

This is also the angle between the initial linear-polarization state $|A_2\rangle$ of the probe and the probe basis state $|w_3\rangle$. Also, the shift δ in polarization between the initial probe states $|A_2\rangle$ and the state $|A_1\rangle$ (the angle between $|A_1\rangle$ and $|A_2\rangle$) is given by

$$\delta=\cos^{-1}\left(\frac{\langle A_1|A_2\rangle}{|A_1||A_2|}\right), \tag{30}$$

or, substituting Eqs. (25) and (28), one obtains

$$\delta=\cos^{-1}(1-4E). \tag{31}$$

6. POVM MEASUREMENT OF ENTANGLING PROBE

Instead of performing a von-Neumann projective measurement of the entangling probe (using the Wollaston prism along with two photodetectors, as in Section 5), one can unambiguously detect the two nonorthogonal probe states $|\alpha_+\rangle$ and $|\alpha_-\rangle$, at least some of the time. For this purpose, the POVM (See Fig. 1 of Ref. 3) will simply be setup to distinguish the non-orthogonal states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ (instead of the states $|u\rangle$ and $|v\rangle$ described in Ref. 3). For this purpose, the Wollaston prism in Fig. 1 of Ref. 3 must be aligned to separate the orthogonal states:

$$\left|\hat{e}_{\alpha_+ +\alpha_-}\right\rangle\equiv\frac{|\alpha_+\rangle/|\alpha_+|+|\alpha_-\rangle/|\alpha_-|}{[(\langle\alpha_+|/|\alpha_+|+\langle\alpha_-|/|\alpha_-|)(|\alpha_+\rangle/|\alpha_+|+|\alpha_-\rangle/|\alpha-|)]^{1/2}} \tag{32}$$

and

$$\left|\hat{e}_{\alpha_+ -\alpha_-}\right\rangle\equiv\frac{|\alpha_+\rangle/|\alpha_+|-|\alpha_-\rangle/|\alpha_-|}{[(\langle\alpha_+|/|\alpha_+|-\langle\alpha_-|/|\alpha_-|)(|\alpha_+\rangle/|\alpha_+|-|\alpha_-\rangle/|\alpha_-|)]^{1/2}} \tag{33}$$

(instead of $\left|\hat{e}_{u+v}\right\rangle$ and $\left|\hat{e}_{u-v}\right\rangle$, as in Ref. 3). The overlap Q between the states $|\alpha_+\rangle$ and $|\alpha_-\rangle$ is given by

$$Q=\frac{\langle\alpha_+|\alpha_-\rangle}{|\alpha_+|\,|\alpha_-|}, \tag{34}$$

or using Eqs. (5) and (6), one obtains

$$Q=\frac{1-3E}{1-E}. \tag{35}$$

The reflection coefficient R_1 of the beamsplitter BS_1 in Fig. 1 of Ref. 3 must, for the case at hand, be given by

$$R_1=\tan^2\left(\frac{1}{2}\cos^{-1}Q\right)=\frac{1-Q}{1+Q}, \tag{36}$$

or substituting Eq. (35) in Eq. (36), one obtains

$$R_1=\frac{E}{1-2E}. \tag{37}$$

Thus the reflection coefficient R_1 must be set, according to Eq. (37), by the set error rate induced by the entangling probe. This will require a beamsplitter with an adjustable reflection coefficient.

The inconclusive rate $R_?$ (or, equivalently, $P_?$ in the notation of Ref. 3, also see Refs. 16–18) of the POVM receiver is given by Ref. 3

$$R_?=\frac{\langle\alpha_+|\alpha_-\rangle}{|\alpha_+|\,|\alpha_-|}. \tag{38}$$

Next, using Eqs. (34) and (35) in Eq. (38), one obtains

$$R_?=\frac{1-3E}{1-E}. \tag{39}$$

The conclusive rate R_c is then given by

$$R_c=1-R_?, \tag{40}$$

or using Eq. (39), one obtains

$$R_c=\frac{2E}{1-E}, \tag{41}$$

which is clearly nonoptimal, particularly for low error rates, as can be seen by comparing Eq. (41) with Eq. (20). However, if the loss in the key distribution channel between the probe and the legitimate receiver, due to attenuation, equals the inconclusive rate, Eq. (39), and only the conclusive states are relayed by the probe to the legitimate receiver, then the entangling probe together with the POVM receiver can obtain complete information on the pre-privacy-amplified key, once the polarization bases are announced in the public channel during reconciliation. Also, to counter alteration in the attenuation due to the probe, the legitimate channel might be replaced by a more transparent one.[2] One may therefore conclude that the BB84 protocol has a vulnerability very similar to the well-known vulnerability of the B92 protocol.

7. SUMMARY

Exploiting the quantum circuit model of quantum computation, the quantum circuit, needed to implement the optimum unitary transformation representing an entangling-probe attacking the BB84 protocol, was determined and shown to yield the correct entangled states. The quantum circuit, faithfully representing the optimum entangling probe, consists of a single CNOT gate in which the control qubit consists of the two photon-polarization basis states of the signal, the target qubit consists of the two probe-photon polarization basis states, and the probe photon is prepared in a specific initial linear-polarization state, set by the induced error rate. The required initial polarization state of the probe photon can be produced by a single-photon source together with a linear polarizer. The gated probe photon, optimally entangled with the signal, enters a Wollaston prism which separates the appropriate correlated states of the probe photon to trigger one or the other of two photodetectors. Basis selection, revealed on the public channel during basis reconciliation in the BB84 protocol, is exploited to correlate photodetector clicks with the signal transmitting the key, and to assign the most likely binary numbers, 1 or 0, such that the information gain by the probe of the key, prior to privacy amplification, is maximal. The probe is a simple special-purpose quantum information processor that will improve the odds for an eavesdropper in gaining access to the pre-privacy-amplified key, as well as impose a potentially severe sacrifice of key bits during privacy amplification.[10]

Finally, it is argued that if the projective measurement of the probe is replaced by a nonoptimal measurement using a POVM receiver, then unambiguous signal state discrimination is achieved at least some of the

time. It follows that if the loss rate due to attenuation in the key distribution channel equals the inconclusive rate of the POVM receiver, then the standard BB84 protocol has a vulnerability very analogous to the well-known vulnerability of the standard B92 protocol.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Laboratory and the Defense Advanced Research Projects Agency. The hospitality of the Isaac Newton Institute for Mathematical Sciences at the University of Cambridge, where part of this work was performed, is gratefully acknowledged.

REFERENCES

1. C. H. Bennett, *Phys. Rev. Lett.* **68**, 3121–3124 (1992).
2. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Rev. Mod. Phys.* **74**, 145–195 (2002). (See p. 152.)
3. H. E. Brandt, *Am. J. Phys.* **67**, 434–439 (1999).
4. H. E. Brandt, US Patent No. 5,999,285 (7 December 1999).
5. C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.
6. H. E. Brandt, *SPIE Proc.* **5436**, 48–64 (2004). [In Eq. (132), $\sin\mu$ and $\cos\mu$ should be interchanged in the coefficient of $|w_3>$ only. Also, $e_{2\theta}$ should be e_θ. In Eq. (195), the overall sign of the coefficient of $|w_2>$ should be $\mp$.]
7. B. A. Slutsky, R. Rao, P. C. Sun, and Y. Fainman, *Phys. Rev. A* **57**, 2383–2398 (1998).
8. C. A. Fuchs and A. Peres, *Phys. Rev. A* **53**, 2038–2045 (1996).
9. H. E. Brandt, *Phys. Rev. A* **66**, 032303 (16), (2002).
10. H. E. Brandt, *J. Math. Phys.* **43**, 4526–4530 (2002).
11. H. E. Brandt, *J. Opt. B* **5**, S557–S560 (2003).
12. H. E. Brandt, *Quantum Inform. Process.* **2**, 37–79 (2003).
13. H. E. Brandt, *Phys. Rev. A* **71**, 042312 (14), (2005).
14. H. E. Brandt, to appear in *J. Mod. Opt.* (2005).
15. H. E. Brandt, Invention Disclosure: *Quantum Cryptographic Entangling Probe* (U.S. Army Research Laboratory, Adelphi, MD, 2004).
16. H. E. Brandt, *Phys. Rev. A* **62**, 042310 (14) (2000).
17. H. E. Brandt, *Phys. Rev. A* **64**, 042316 (5) (2001).
18. H. E. Brandt, *Contemporary Math.* **305**, 43–52 (2002).

Quantum Information Processing, Vol. 4, No. 6, December 2005 (© 2006)
DOI: 10.1007/s11128-005-0007-9

Multiple RF Coil Nuclear Magnetic Resonance Quantum Computing

Lisa C. Siskind,[1] Bruce E. Hammer,[1,4] Nelson L. Christensen,[2] and Jeffrey Yepez[3]

Received July 7, 2005; accepted December 5, 2005; published online February 24, 2006

Recent work has demonstrated the feasibility of using an array of quantum information processors connected via classical channels (type II quantum computer) to implement a quantum lattice-gas algorithm. This paper describes work towards constructing a new experimental set-up for a type II quantum computer. This set-up has new hardware and software specifications but does follow previously published approaches of operation encoding the initial mass density onto a twoqubit processor and using standard pulse techniques to step through the algorithm. New hardware for this system includes the ability to read both qubits at once, effectively reducing the processing time by twofold. Hardware changes also include the use of multiple coils controlled by a single spectrometer and a hardware switch. New software includes a top level control system for the spectrometer for quick experimental configuration as well as configurable modeling software to verify results. Results are presented here from a system with the final software implementations and the two channel spectrometer configuration run on a single prototype coil. Progress towards the final multi-coil implementation is described.

KEY WORDS: Quantum information processing; nuclear magnetic resonance; quantum lattice gas; diffusion equation; quantum computing.

PACS: 03.67.Lx; 82.56.-b.

1. INTRODUCTION

Researchers in the newly emerging field of quantum information processing have made great strides both in its theoretical development and experimental practice during the last several years but progress is still strongly

[1] Center for Interdisciplinary Applications in Magnetic Resonance, University of Minnesota, Minneapolis, MN 55455, USA.
[2] Carleton College, Northfield, MN 55057, USA.
[3] Air Force Research Laboratory, Hanscom Field, MA 01731, USA.
[4] To whom correspondence should be addressed. E-mail: hammer@umn.edu

limited by the technological difficulty of accurately preparing and controlling the quantum state of many qubits. One successful approach to expanding the computational capabilities of current quantum information processing hardware is the construction of type II quantum computers. Type II quantum computers, as described in Refs. 1,2, constitute an ordered set of quantum processing nodes where inter-nodal information transfers is reducible to invertible quantum mechanical operators represented by orthogonal matrices – effectively inter-nodal communication via classical channels. This architectural paradigm allows for the use of several, uncorrelated quantum processors to speed computational times as a whole by working on separate parts of the problem: quantum parallelism occurs locally within a quantum node while classical parallelism occurs globally. A group from the Massachusetts Institute of Technology (MIT) has implemented one such device using a liquid state nuclear magnetic resonance (NMR) imaging technique and have achieved proof-of-concept by simulating the spacetime dynamics of a quantum lattice gas in 1+1 dimensions, modeling the macroscopic scale behavior of a many-body quantum system whose effective field theory is either the classical diffusion equation[3–5] or the classical nonlinear Burgers equation.[6,7] This type II system used a standard NMR setup with field gradients to spread out the magnetization signal from a single sample, containing two spin-1/2 nuclei, into 16 effective quantum computational nodes. While offering great advantages in increasing the number of computational nodes available, it also had the disadvantage of the complexity of the operations needed to encode and readout out data while still preserving the distinctness of the quantum nodes.[5]

This report describes our work towards implementing a different design for a type II NMR quantum computer intended to address the issues mentioned above.[6] Furthermore, this report presents new experimental results

---

[5]A newer version of the MIT approach attempts to employ a k-space encoding that directly implements the inter-nodal communication (or stream operator) using pulsed field gradients. This eliminates the need for continual readout of data after each time step.

[6]At University of Minnesota Center for Magnetic Resonance Research, a modern commercially available magnetic resonance imager called SENSE is used, arrays of four or eight different coils and receivers have been commonplace for the last five years. In some ways, they are built with vastly more sophistication than the idealized prototype system we have considered. Clinical SENSE coil arrays are a series of overlapping surface coils used to limit penetration and increase the signal to noise for imaging purposes. They are inappropriate for QIP applications for the following reasons. First, NMR QIP uses different nuclear spins to serve as qubits. This requires the use of multiple tuned or multiple coil NMR probe. Multi-frequency SENSE arrays are not presently available and will take a considerable engineering effort to develop. Second, QIP requires a highly uniform B_1 field for QIP calculations. SENSE coils have an inhomogeneous B_1 field due to their planar structure. Adiabatic pulses are not appropriate for NMR QIP because it takes milliseconds to achieve

as well as new NMR simulator results (modeling the operation and error sources in the NMR spectrometer) of the same quantum lattice-gas algorithm tested in Ref. 5. The main goals of this work are to explore this new hardware approach, improving efficiencies where possible, and streamlining and improving related data acquisition and control software as well as the simulator software. This work is the first step in an effort to construct a practical multi-coil system that can be quickly reconfigured and reused for testing different quantum computing algorithms.

Two key differences exist between our hardware design and the MIT design. The first is the simultaneous use of two receive channels as well as two output channels. This is in contrast to most system configurations that allow one or more send channels but are restricted to a single receive channel. This system can be scaled for additional send and receive channels as well. The advantage of simultaneous data readouts for each coil is twofold: the first is in performance and the second in accuracy. The system performance speed is doubled since the data can be obtained from two channels at once rather than having to read one channel, repeat the experiment, transfer the data from one channel to another and then read out the second channel. Furthermore, the experimental accuracy also improves since the data from both channels are obtained from the same run and not a repeat of the previous run; hence we can avoid errors from irreproducibility problems.

The second difference is using a one-to-one mapping of NMR coils to quantum computers as depicted in Fig. 1. This means as many coils are needed as desired quantum processors. This offers the advantage of relatively simple encode and read operations to operate each processor. It also allows the system to let nuclear spins contained within individual coils relax in preparation for the next computation while another coil is

---

a spin rotation versus microseconds in a conventional high homogeneity RF coil. Third, SENSE coils require precise geometric layouts. Crosstalk between small dimension SENSE coil arrays will pose significant, if not insurmountable, engineering challenges. They isolate themselves from one another by careful geometric arrangement of the coil layout. As the dimension of the coil decreases, the relative error in layout increases – thus making orthogonality of RF fields from adjacent coils more difficult. In contrast, in our approach the exact location of each microcoil, with respect to one another, is not critical. This is because the design goal is to have each microcoil be self-shielded, hence mitigating inter-nodal interaction. Positional uncertainty is not an issue in our case as it is with SENSE coils. Forth, typical SENSE coils have 20 dB isolation between channels. Under extraordinary conditions SENSE coils can have up to 60 dB isolation from one another. Crosstalk between channels will lead to decoherence in QIP experiments. Using individually shielded RF coils can provide isolation exceeding 100 dB thus providing higher immunity from decoherence. Finally, QIP uses ^{1}H, ^{13}C and other spin-labeled compounds to serve as qubits. A minimum of two qubits (different spins) are necessary for liquid sample QIP applications. Expensive "designer" molecules are necessary for two qubit applications, e.g., $^{13}CHCl_3$. Samples for QIP applications are mass limited and cannot use available SENSE coils due to filling factor and SNR considerations.

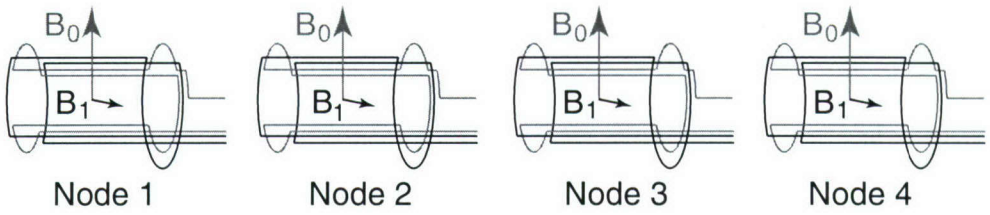


Fig. 1. Each coil acts as a quantum processor.

run by the central control processor and spectrometer, a type of pipelining. This is especially handy for samples with long T_1 times. Finally, this setup also allows for optimal use of costly spectrometer channels, increasing the system speed with the addition of more relatively inexpensive coils.

The MIT approach uses pulsed field gradients, decoupling and shaped radio frequency (RF) pulses to subdivide one sample into multiple quantum processors. It is difficult to create isolated quantum processing cells by this method. There is the potential for cross-talk between adjacent cells created by shaped RF pulses and gradients. The RF pulses are typically amplitude and phase modulated to create well defined spatial nodes. However, the dividing lines between the spatial nodes are not "brick walls" and thus the spins in adjacent slices are not completely isolated.

The use of shaped pulses also leads to a dispersion of phases within the selected slice. For many collision experiments there can be a significant accumulation of phase errors. The imaging approach also requires decoupling during the read portion of the experiment. This leads to partial spin polarization due to the nuclear overhauser effect (NOE).[8] NOE increase in signal strength for a carbon spin can be up to a threefold factor during proton irradiation and depends on the efficiency and duration of decoupling. This is an additional source of error when integrating the detected NMR signal.

A pulsed field gradient (PFG) is necessary to spatially map the spins. A PFG will induce eddy currents within the bore of an NMR magnet resulting in a time-dependent perturbation of the spins lasting long after the gradient is off. This can result in imperfect slices and unwanted frequency and phase shifts of the NMR spin ensemble. Note that PFGs are necessary for destroying off diagonal spin coherences during an NMR quantum information processing experiment and is a major source of error for both the imaging and multiple coil approach.

Finally, subdividing a sample by imaging methods results in a smaller signal to noise ratio (SNR) for the defined voxel than attainable by having a coil with a dimension equivalent to the voxel size. This observation is known as reciprocity and is the basis of matching the RF coil dimension to the sample size to optimize SNR.[9,10] The SNR for a specific nucleus at a fixed temperature is defined by

$$\frac{S}{N} \propto \frac{B_1}{I} \frac{V_s N \omega_o^2}{V_{noise}}, \tag{1}$$

where B_1/I is the RF coil sensitivity, V_s the sample volume, N the number of spins, ω_o is the larmor frequency and V_{noise} is the sample and RF coil noise voltage.[9,11] The term B_1/I represents a unit current induced in an RF coil from nuclear spins precessing in the transverse plane. This term becomes larger when spins coupling to the coil are physically close to the coil. So as the coil dimension decreases, nuclear spins are closer to the coil windings, thus inducing a larger current in the coil. The behavior of RF coils for mass-limited samples, i.e., a fixed number of spins, has been theoretically and experimentally investigated.[9,12] Generally, for a solenoidal RF coil the SNR per unit volume of sample is proportional to the inverse of the coil diameter for a coil diameter greater than 0.1 mm. For coils below 0.1 mm in diameter the per unit volume SNR is proportional to the square root of the coil diameter.

In addition to the new hardware features mentioned above, several custom software applications were created. The first was a highly reconfigurable control program that runs on the top layer of the experiment, controlling both the spectrometer software and coordinating the data collection and writing. This software was designed and written in a modular form and can be quickly reconfigured for different algorithms and experiments. The second software application was an NMR simulator to verify experimental results. This consisted of a Java user interface and a mathematica back-end. Again, both pieces were designed in a modular fashion for ease of reconfiguration and reuse.

For the initial work detailed in this paper, a single prototype hardware coil was used to simulate the behavior of 16 parallel coils. The software written for this NMR quantum computer experiment was designed for use with the multiple coil system.

This paper will first present a brief description of the quantum lattice gas system in Sec. 2. Section 3 describes the NMR implementation hardware and software. Section 4 shows the results of the first data run with the prototype coil and Section 5 provides a short summary and remarks about future work.

2. QUANTUM LATTICE GAS SYSTEM

The type II quantum algorithm can be exactly represented at the mesoscopic scale by an effective field theory: the *quantum Boltzmann equation* of motion that describes the time-dependent dynamics of a kinetic many-body particle system. The quantum algorithm is a novel computational

technique to reduce the complexity of fluid dynamics simulations where quantum logic operations directly causing qubit–qubit interaction are used to model the relevant local particle–particle interactions occurring within the fluid. The quantum model of the diffusion equation, in $D+1$ dimensions where D can be any number of spatial dimensions, requires only two qubits per node, and hence encodes the occupation of up to two particles at that node. An operator splitting method reduces the algorithmic complexity down to a modeling problem in $1+1$ dimensions.[13] This decomposition method in demonstrated for a $2+1$ dimensional test case in Ref. 14. The quantum algorithm then models unit-mass particle–particle interaction with one unitary collisional operation per lattice node per unit time step. In particular to model diffusion in $1+1$ dimensions, the $\sqrt{\text{SWAP}}$-quantum logic gate models the on-site particle–particle interaction. Once the interaction calculation is completed, the final algorithmic step of a single iteration of the time-dependent evolution of the quantum algorithm requires streaming the results (particle occupations) to the neighboring nodes. Both the collisional operation and the streaming operation conserve mass over the entire system. The orthogonal stream operator, $\mathcal{S}$, and its transpose (inverse), $\mathcal{S}^{\mathrm{T}}$, are the two informational transfer operations on the one-dimensional lattice, one operation per direction ($\pm x$). In the case where the particles have an equal probability of reversing direction in each interaction, the particle motion is effectively a random walk and the mass density field consequently diffuses isotropically over time.[5,15]

Our work concentrates on the implementation of the type II quantum algorithm in $1+1$ dimensions. This algorithm works by mapping the energy of each particle to the probability values of our quantum system and uses quantum state mixing to describe the interaction of particles.[5,16,17] Mapping the occupational probability onto the single-particle state of a quantum bit is done as follows:

$$|q\rangle = \sqrt{f}|1\rangle + \sqrt{(1-f)}|0\rangle. \tag{2}$$

A two qubit system is required for a one-dimensional quantum lattice representation of the diffusion equation. The initial wave function for such a system with f_1 and f_2 occupational probabilities is the following tensor product state:

$$\begin{aligned}|\psi(m,n)\rangle = {} & \sqrt{f_1(m,n)f_2(m,n)}|11\rangle + \sqrt{f_1(m,n)(1-f_2(m,n))}|10\rangle \\ & + \sqrt{(1-f_1(m,n))f_2(m,n)}|01\rangle \\ & + \sqrt{(1-f_1(m,n))(1-f_2(m,n))}|00\rangle,\end{aligned} \tag{3}$$

where m is the node index and n is the time-step index.

The four main steps of the type II quantum lattice gas algorithm are:

1. Encoding of the probability value for each qubit (i.e., f_1 and f_2).
2. Application of a unitary collision operator for quantum state mixing:

$$|\psi'(m,n)\rangle = \hat{U}|\psi(m,n)\rangle, \tag{4}$$

for all m where the $\sqrt{\text{SWAP}}$-quantum logic gate is

$$\hat{U} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2}-\frac{i}{2} & \frac{1}{2}+\frac{i}{2} & 0 \\ 0 & \frac{1}{2}+\frac{i}{2} & \frac{1}{2}-\frac{i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

3. Reading of the resultant probability values f_1' and f_2' (i.e., quantum state reduction).
4. Unitary streaming the results to the next site:

$$f_1(m,n+1) = f_1'(m-1,n), \tag{6}$$

$$f_2(m,n+1) = f_2'(m+1,n). \tag{7}$$

The reading method employed in this experiment follows the method of.[5]

3. NUCLEAR MAGNETIC RESONANCE IMPLEMENTATION

Liquid-state NMR was used for this implementation using a two gram sample of Carbon 13 labeled chloroform ($^{13}CHCl_3$) where the two spin-1/2 nuclei of hydrogen and the labeled carbon 13 comprised a two qubit system for ensemble quantum computing. A frequency difference, or *chemical shift*, of the spin-1/2 nuclei in the labeled chloroform is caused by the local magnetic fields within the molecule's geometrical structure. An external magnetic field B_0, on the order of a Tesla, creates a difference in population of the quantum states, between the aligned and anti-aligned spin-1/2 nuclei because of a small energy change in the Boltzmann weights at thermal equilibrium.[18] This energy difference is relatively small at room temperature but is still detectable due to the large number of molecules in the sample. Spin–spin coupling, mediated by correlated electrons within the molecule, further splits the ground state energy of one of the fermionic nuclei in the molecule depending on whether the other fermionic nuclei is spin up or spin down. This energy shift is much

smaller than the chemical shift, but is also clearly detectable. NMR spectrometers use high fidelity RF circuits to manipulate the net magnetization of the system and detect the population of each state as it returns to equilibrium. For the two qubit sample used above, the system is accurately described by the density matrix that depends on only the intramolecular spin degrees of freedom by tracing over all non-spin degrees of freedom, such as positional degrees of freedom, leaving a reduced 4×4 density matrix.

While NMR is technically advanced as compared to many quantum computing schemes due to the near 60-year history since the publication of the phenomenon of magnetic induction[19] there are still several technical challenges to the use of liquid-state NMR for quantum computing. Inhomogeneities in the background magnetic field and the RF coils cause spin–spin decoherence, limiting the number of quantum logic operational that can be executed with phase-coherence. Another problem is the relatively small number of controllable qubits within a molecule. Yet, the largest problem is the control of the pure quantum state of each molecule individually, which is not possible within the context of our experimental approach; hence the system is always in a mixed quantum state. Liquid-state NMR quantum computing was made possible with the development of several methods[20,21] for ensemble qubit preparation and the creation of start states that simulate pure quantum states. This work used the method described by Price *et al.*[22] for initial state preparation.

3.1. Mapping the Quantum Lattice Gas to the NMR Spin System

Much of this work, but not all, followed the same mapping scheme for the quantum state description to the physical experiment as in previous work.[5] One exception is the elimination of the need for the overall pulsed field gradient and a truncated RF modulation function for the discernment of different quantum processors. A summary of all the implementation differences are listed in Table 1.

3.2. Hardware

The basic elements of a NMR spectrometer setup are shown in Fig. 2. Components of the system include the magnet that provides the external magnetic field, the coil creating the orthogonal magnetic fields to manipulate the sample, the sample itself, a RF signal source that creates the pulse signals, a RF transmitter/receiver to prepare the outbound pulse and receive the resultant signal from the sample and finally the control computer for the whole system. In addition to the main components

Table 1. Implementation differences for type II quantum computer using imaging techniques versus separate coil design

Item	Pravia *et al* (5,16,17)	Multi-coil Implementation
Quantum Computer	Portion of a sample in a coil	One coil per quantum computer
Addressing	Via field gradient *possible crosstalk, possible eddy current artifacts*	By coil *no crosstalk, no eddy current artifacts*
Encoding	Shaped pulses *Decoupling needed*	Hard pulses *No Decoupling needed*
Reading	*Decoupling needed*	*No Decoupling needed*

needed, there are several key items for increasing the quality of the signal. Auxiliary magnetic shim coils are used within the magnet itself for increasing field homogeneity and for the creation of magnetic gradients. Pre-amps are used for increasing signal strength from the mixer and Transmit-Receive (TR) switches that allow the same coil windings to do dual duty as senders and receivers of signal from the sample.

A 1.5 Tesla large bore (55 cm) magnet was used for this experiment. The large bore allowed for the quick prototyping of a custom coil out of readily available materials. This bore can easily accommodate the addition of many more coils for the future expansion of this system. This is the planned next step in this work with the final goal being the construction of at least four coils working in parallel. The prototype coil used orthogonal Helmholtz-saddle circuits to independently control of each spin-1/2 ensemble via RF pulses.

The system described here differs from previous implementations of this algorithm[5,23] in that the entire RF coil is used as a single quantum computer. Previous implementations used a pulsed field gradient to split

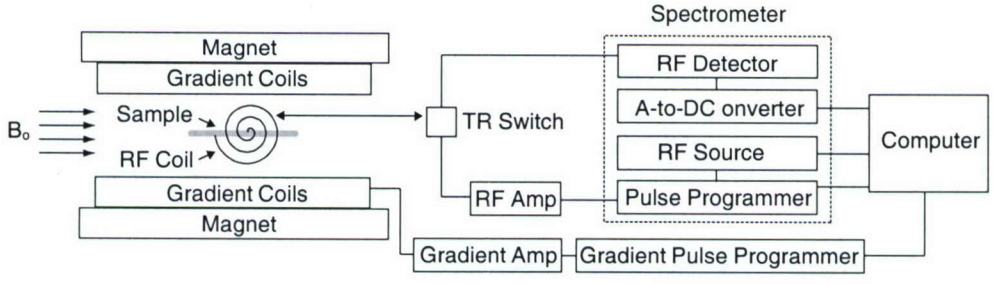


Fig. 2. Schematic overview of a NMR system.

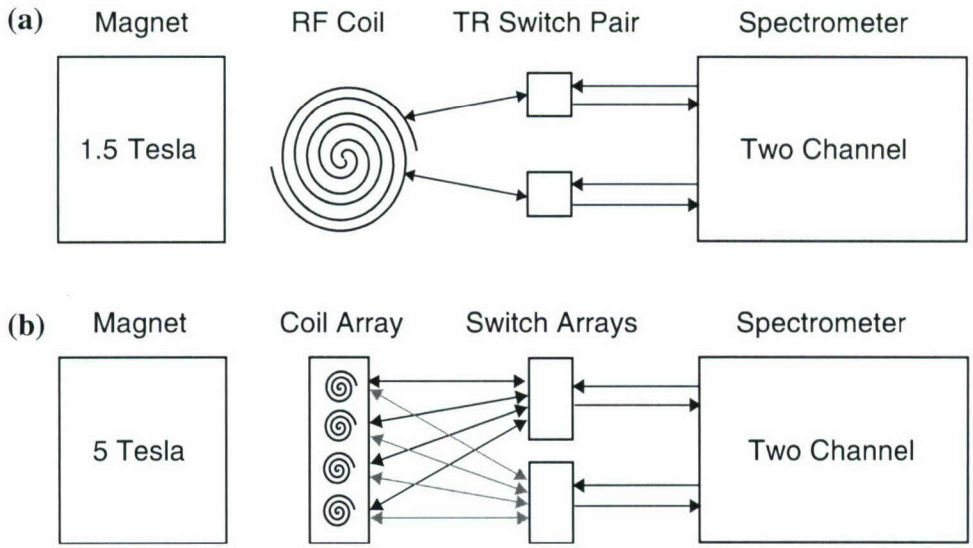


Fig. 3. (a) The prototype configuration shown above was used to obtain the data described in this work. (b) The new setup uses a four coil array, as shown in the final configuration schematic. The four coils are connected to a single multichannel spectrometer via a hardware switch. This switch allows the use of a single coil for a calculation while the remaining coils relax. This significantly improves the use factor for the spectrometer and magnet hardware since the relaxation time used between calculations for this work was 180 s, while the calculation sequence itself is executed in under a second. The use of four coils accessed round-robin instead of repeatedly using a single coil would allow a calculation to be executed every 45 s instead of once every 3 min, decreasing the overall experiment time needed by a factor of four.

a single coil into 16 effective quantum computational processors using NMR imaging techniques. While the single computer per coil design does force the construction of many coils versus the use of a single one, it also allows to us use hard pulses rather than shaped pulses for data encoding and to skip the use of decoupling sequences altogether for encoding and reading channels. It further eliminates artifacts from eddy currents induced by pulsed field gradients and possible crosstalk between adjacent cells due to imperfect slice selection from a truncated RF sinc function all through the use of independent coils.

The eventual goal of this work is to have several such coil quantum processors connected to the spectrometer via a hardware switch as shown in Fig. 3. This would allow the spectrometer to access one coil for a calculation while the remaining coils relax. This approach would significantly improves the use factor for the spectrometer and magnet hardware since the relaxation time used between calculations for this work was 180 s, while the calculation sequence itself is executed in under a second. The use of four coils accessed round-robin instead of repeatedly using a single coil would allow a calculation to be executed every 45 s instead of once every

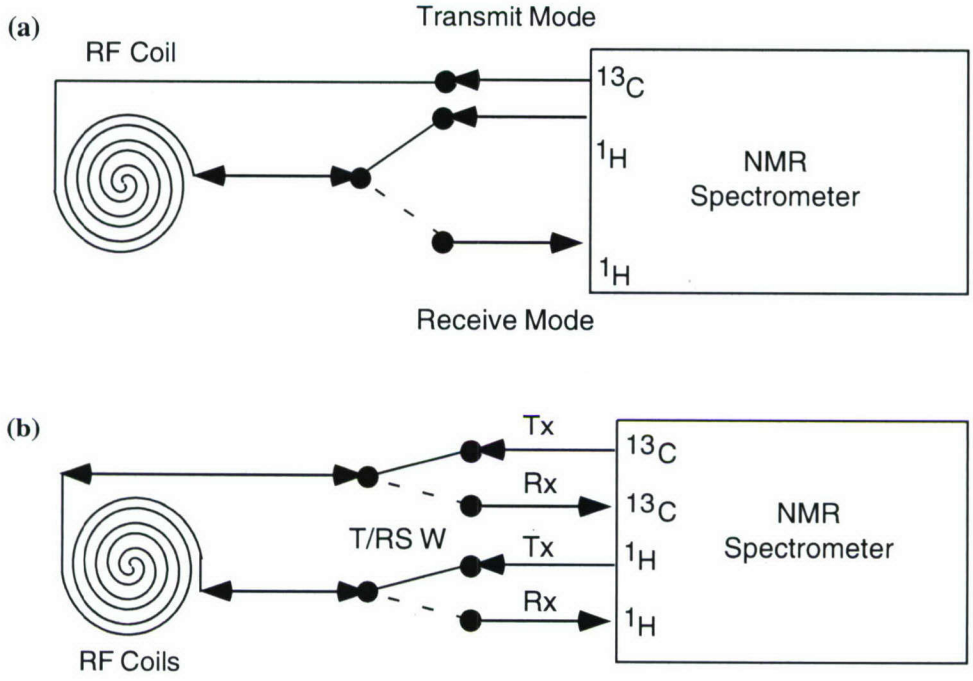


Fig. 4. (a) Classic spectrometer design with two send channels and a single receive channel. (b) Full two channel design configuration used for this work, containing two send as well as two receive channels.

3 min. This decreases the overall experiment time needed by a factor of four while still maintaining all the benefits of long decoherence times.

This approach is feasible within the physical parameters of existing NMR hardware. The magnet allocated for this final work is 210 mm useable bore 5.0 Tesla magnet with uniformity of 2 ppm for a 100 mm diameter spherical volume (DSV). Four individual 22 mm coils will be placed within in an array with center axis separation of roughly 45 mm, easily fitting within the 100 mm DSV. Shim and pulse power calibration parameters can be measured per coil and used by the software control program to minimize response differences between the coils. Additionally, the sample sizes used for these coils sizes are not much larger than those used for micro coils,[9,10] the use of which would allow even more coils to fit within the magnet's 100 mm DSV.

A custom built Apollo spectrometer (Tecmag, Inc., Houston, TX) was used to control the hardware system. This windows based unit was outfitted with two send as well as two receive cards which allowed for simultaneous control of both channels in all aspects of the experiment. This is shown schematically in the bottom section of Fig. 4. This system differs from the typical NMR spectrometer setup in that while many have multiple channels to send pulses out, few have more than one hardware channel

available for reading data. Thus most systems cannot look at the state of more than one qubit at the end of a computation. For a two qubit system this limitation would force an experimenter to run the same computation twice: once for the nominal qubit the channel was tuned to and a second time to read the data from the other qubit, using a swap operation to move the data over to the read-capable channel.

3.3. Control and Simulation Software Design

Custom software was written for both the control of the NMR quantum computer and the simulation software to verify the experimental results. Figures 5 and 6 show the designs for the experimental and the simulation software, respectively. A high level of encapsulation was used in the design for both systems. This allows for maximum reusability for the core code, meaning that core code written for this experiment could also be used for future experiments, whether a different algorithm for liquid state NMR quantum computing or even the same algorithm for solid state quantum computing.

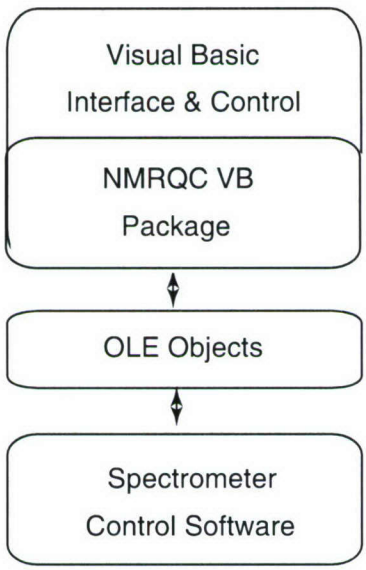


Fig. 5. High level architecture of the Experimental Control Software. This software is comprised of several layers: the top layer user interface and control module, the OLE objects used to communicate to the spectrometer and the spectrometer software itself. The top layer is designed to eventually be split into two separate layers as is denoted by the second object contained within it labeled "NMRQC VB Package". This package contains core NMR quantum computing commands needed for the execution of any quantum algorithm.

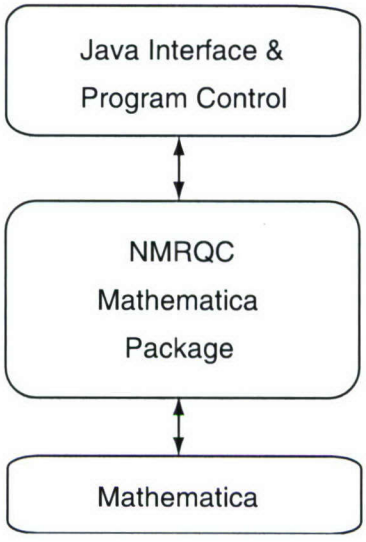


Fig. 6. Design of the Simulation Software. This software is comprised of several layers. The top layer contains the user interface and control module, all written in Java. The middle layer is the package of generic quantum computing Mathematica commands needed for the actual calculations. The middle layer also contains a few helper commands that combine many of the generic functions for quick communication. The bottom layer is the Mathematica engine itself that executed the commands and returns the results. Not explicitly shown is the J/Link communication package that facilitates communication between the Java application and the Mathematica software.

3.3.1. Experimental Control

The experiment was run using a Visual Basic software program developed to execute the algorithm logic and work as a higher layer of control over the spectrometer software. This allowed for precise timing between sequence executions as the sample reset to the thermal state. It also allowed for the automated analysis of results.

The Visual Basic program consisted of two layers: "User Interface and Control" and a core package of commands, "NMRQC VB Package", to access the Tecmag Objecting Linking and Embedding (OLE) objects[7] which were in turn used to control the spectrometer software and hardware.

The "User Interface and Control" layer accepted user input for experiment start values, provided a method to control the start of the experiment, coordinated the execution of all core commands needed to implement the algorithm and provided user feedback on the progress of an experiment underway. This allowed for the running of multiple

[7]OLE is a Microsoft developed standard for access and control of one program by another. The Tecmag software is written following the OLE standard and the company provides an Application Program Interface (API) guide for accessing these objects on their system.

experiments with different initial mass density distributions without the modification of a single piece of code.

The control layer also handled the saving and initial analysis of data. Data was stored as both raw spectrum files as well as final integrated values. The raw spectrum showed the state of the system over the duration of the experiment, as well as allowed the possibility of later applying multiple analysis methods, if so desired.

The layer just below the user interface, the "NMRQC VB Package", contained general commands needed for NMR quantum computing. This core code contained commands that are needed to run NMR quantum computing experiments in general, such as operations to initialize the system, encode values into each qubit and read out resulting values.

3.3.2. Pre-computed Experimental Overhead

In an ideal world, the sample response to RF power would be linear and could accurately be characterized by percent power needed per degree flip angle, which is in turn related to the spectral integral projection value by a sine function. This was not the case for the actual hardware in the real world. Others have found this to be true as well[16] and used a spy coil to record the actual response and adjust accordingly. The system used for this experiment was consistent enough to allow the use of a lookup table for actual found responses per RF power unit. The creation of this entailed the following steps:

1. The nominal RF power needed to flip the sample spectra 180 degrees in 180 or 240 μs was found, depending on the system used.
2. The RF power was divided into 28 equal parts.
3. The overall time used for the RF pulse was reduced by a factor of two so the nominal 180 power should produce a 90 degree pulse.
4. Spectra integral values were measured for the 28 RF pulse intervals chosen above plus another additional division in the event of undershooting the initial 180 degree power for a total of 29 RF power intervals.

These steps were repeated for each channel used. This created a lookup table of percent maximum response per percent input power. For the running of the actual experiment needed RF pulse powers were interpolated between lookup table values as needed. Figure 7 shows the data for the 1.5 Tesla system.

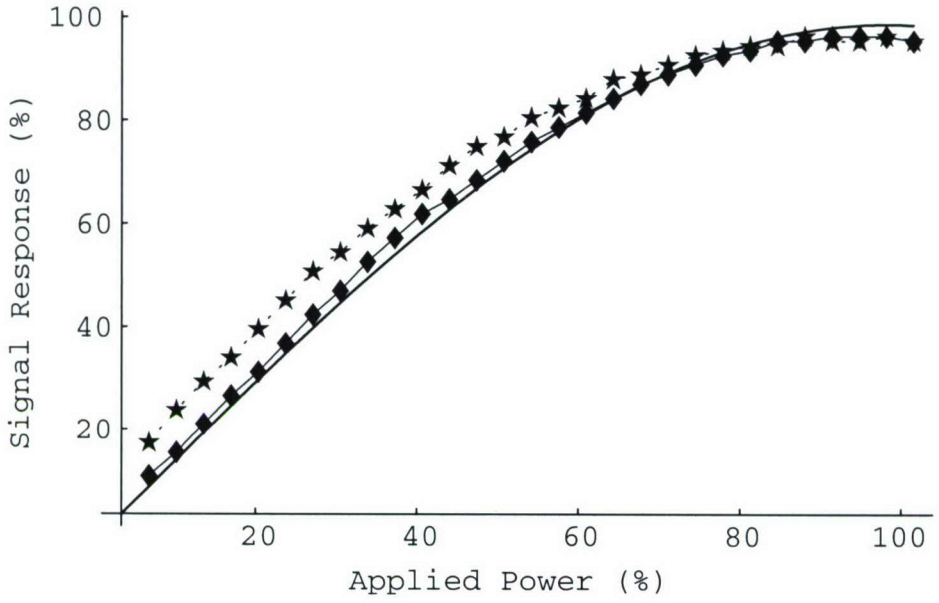


Fig. 7. The percent of maximum signal response, per percent input RF power from the spectrometer software. Twenty nine divisions were used to measure the overall system signal response. These are the maximum power divided by 28, plus another increment to verify that the maximum power was not under-estimated. The maximum power value was determined during an initial calibration process as the power needed to produce a 90 degree flip, i.e., the maximum signal. The diamonds are channel one and the stars are channel two. A sine curve, showing the theoretical power response, is fitted to channel one and is shown as the solid line.

3.3.3. Simulation of Expected Results

Experimental results were verified through a separate simulation program. This program also follows the multilayer paradigm, as shown in Fig. 6, for a high level of code reuse. This software consists of two main sections: a core Mathematica package containing all the quantum state manipulations necessary and a top-layer java application written to access the Mathematica engine through the Mathmatica toolkit J/Link. This design allows the top layer java application to evaluate different user input, load the core quantum computing package, execute the desired simulation and quickly return well formated results. The benefits of such a design are multi-fold: the top layer user interface allows different inputs to be evaluated nearly instantaneously without any re-coding, standardized graphical user interface tools can be used, along with standardized logging techniques for the results, and finally, the lower layer core Mathematica package can be reused for new quantum algorithm experiments.

The simulator program calculated two sets of data. The first is an ideal calculation dataset that assumes perfect encoded values and zero

error terms, for each cell of our simulated lattice. The second dataset evaluates the addition of user specified errors in the implementation, such as incorrect system calibration and outside noise. The error terms specification allowed for the evaluation of both systematic as well as random errors. The systematic errors analyzed included incorrect flip angle power calibration, which would introduce overall under or overshooting of the intended angle, incorrect initial value encoding error, which is also a flip operation but was analyzed separately, incorrect evolution timing, overall signal loss and incomplete gradient applications. Random errors analyzed included random flip angle errors, evolution errors and encoding errors.

Systematic errors were modeled by using a multiplier of magnitude selected by the user. It could be positive or negative. All operations using the selected operation were equally affected. Selecting an error of 0 is the same as executing the ideal calculation. Furthermore, random errors were also propagated by a multiplier but this time using a random number generator to selected a number within plus or minus the user selected value. A new random number was selected for each occurrence of the operation, and for each iteration.

Overall loss was calculated by reducing the final output matrix values by the user selected amount at the end of each complete calculation. This was performed once per calculation cell per iteration.

Incomplete gradient application was calculated by attenuating the density matrix off-axis values to a user selected percentage. Ideally, the application of a gradient reduces all off-axis terms of the density matrix to zero. The result of this application was the existence of off-axis terms after the application of a gradient. This error was termed gradient "leakage," since the magnitude denoted the percent of off-axis terms remaining.

3.4. Quantum Lattice Gas Simulation

Figure 8 shows the steps involved in one calculation cycle. The first portion manipulates the state density into the pseudo-pure start state.[22,24,25] This is a much simpler start state then the nominal room temperature mixed state, both equalizing the Hydrogen and Carbon signals and eliminating off-axis and mixed states. Once this is accomplished, the system is encoded with the desired start values. Both qubits are encoded at the same time followed by the application of the collision operator. Finally, the values of both qubits are read.

Using hard pulses and separate coils for each calculation rather than imaging techniques to divide the sample into effective cells in frequency space eliminates the need for decoupling during encoding and reading of data. A small spread in frequency space does not impact the accuracy of the separate coil implementation since it does not use that domain to define the

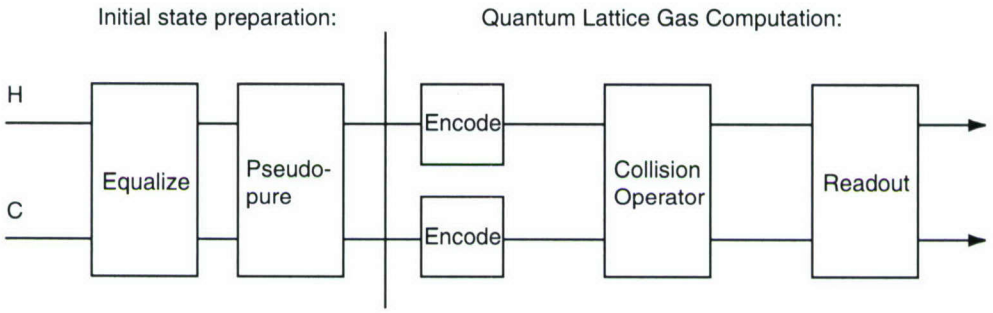


Fig. 8. The pulse sequence of single time step for the NMR implementation of the quantum lattice-gas algorithm. The first steps prepare the sample for a calculation by equalizing the spin signals and then setting them into a pseudo-pure state. Once this is done, each spin ensemble is encoded with the desired starting values. The collision operator is applied to the sample and final values of the resultant density values are ready out. Both channels are read at the same time. Hard pulses are used throughout.

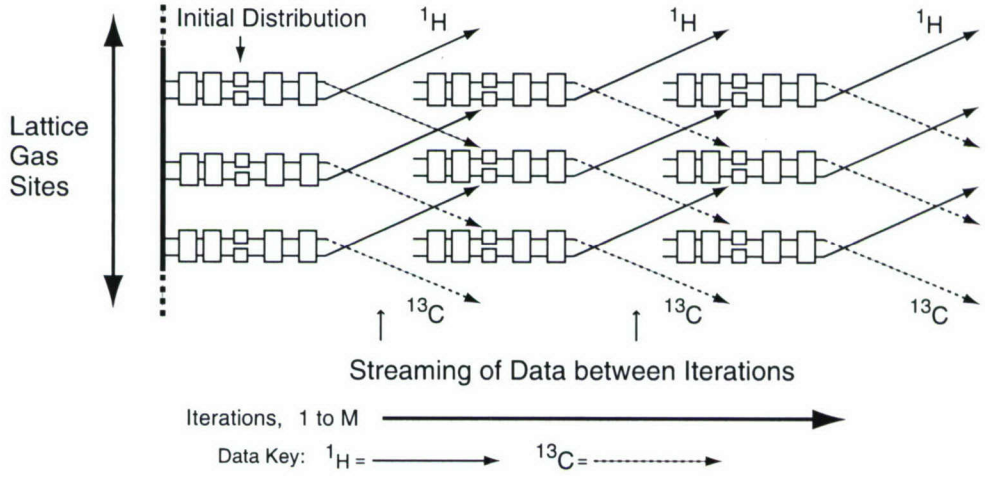


Fig. 9. The data flow for the lattice-gas algorithm, adapted from.[16] The calculations for the algorithm were run one at a time on the single prototype coil, simulating the action of 16 different coils running simultaneously for each iteration of the algorithm. See Fig. 8 for details on each quantum circuit. The read values for each calculation were streamed to the next encode step at the appropriate location in the lattice for the next iteration: the hydrogen moving up one step and the carbon moving down one step between each iteration.

calculation cells themselves. This saves in both execution time and accuracy. Both channels were encoding during a single 90 μs pulse time step rather than the two 8 ms pulse intervals needed to encode the data for the previous implementation using imaging techniques.[5] Additionally, exposure to well know decoupling error terms such as the NOE are eliminated.

Figure 9 shows the flow of data over the course of the entire experiment. Values are streamed between iterations. A continuous boundary

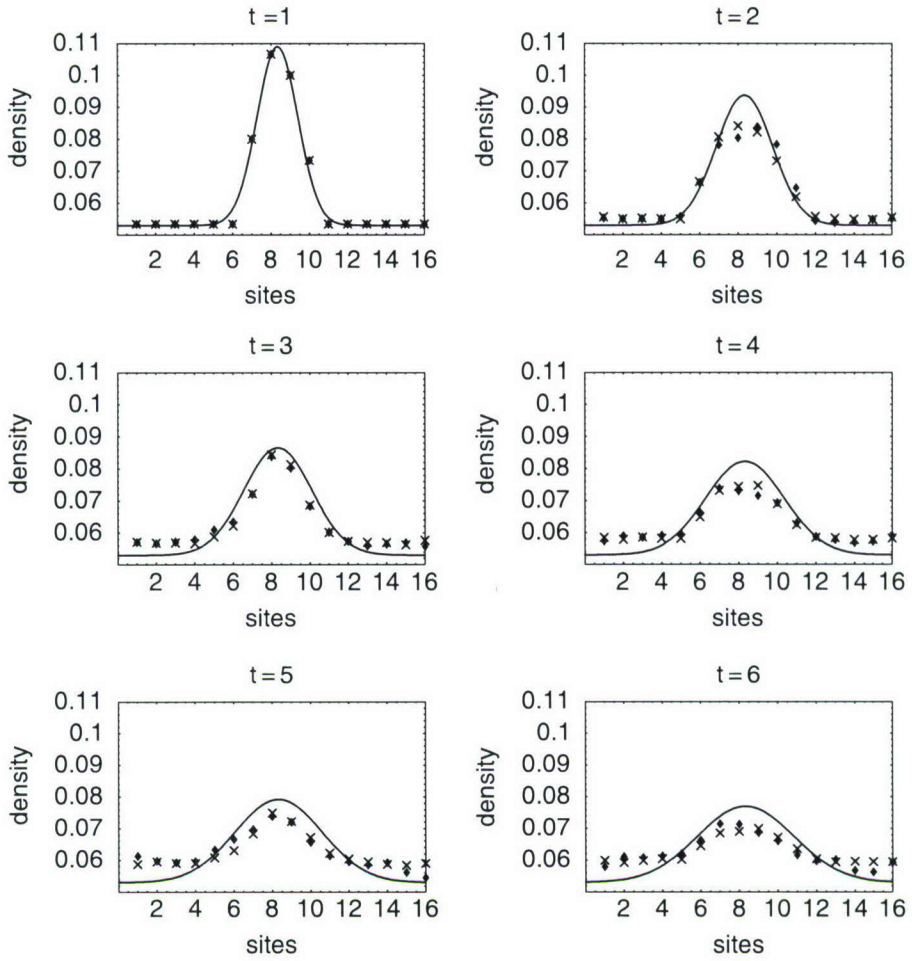


Fig. 10. Ideal (solid curve), experimental (diamonds) and simulated (crosses) values for the 16 cells of the quantum lattice-gas algorithm over six time steps. The algorithmic time labels each snapshot. For each graph, cells number one through 16 are shown along the bottom from left to right. The cell values of the total number density are shown vertically.

condition is met by having the values of quantum computation cell 16 streamed to cell 1 and vice versa.

4. EXPERIMENTAL RESULTS

Figure 10 shows the ideal, actual and simulated data for the 1.5 Tesla system. There is good over all agreement between the analytical solution and experimental values in terms of the overall shape, but there is an overall growth in the density values across all cells for each iteration of the algorithm.

The experimental data was simulated with reasonable agreement using the error values shown in Table 2. Analysis of the data presented

Table 2. Simulator error values used to model the experimental data

Item	Systematic (%)	Random (+/−%)
Error values used for simulated data		
Evolution	−2.0	0.5
Flip	3.5	0.5
Proton encode	−4.0	0.5
Carbon encode	−2.0	0.5
Overall loss	1.0	0.0
Gradient leakage	35	n/a

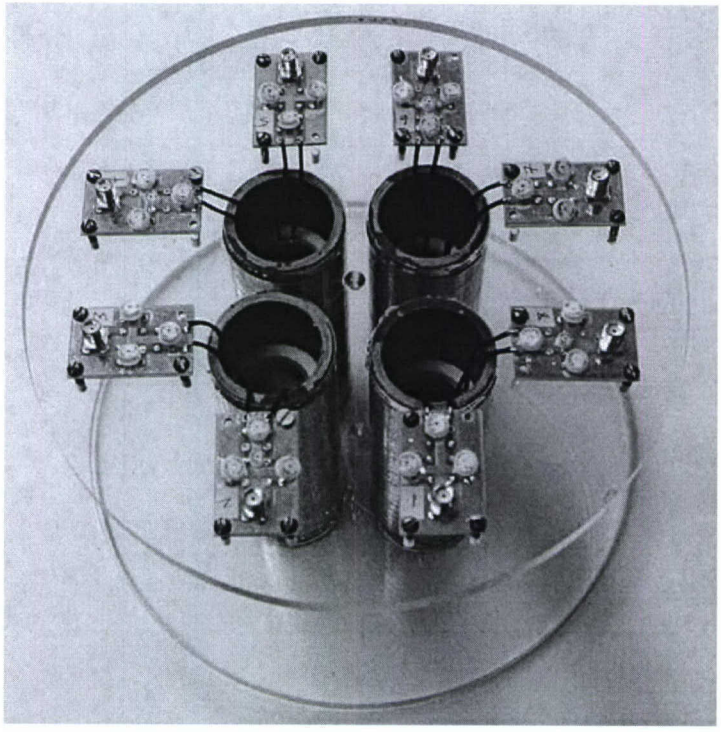
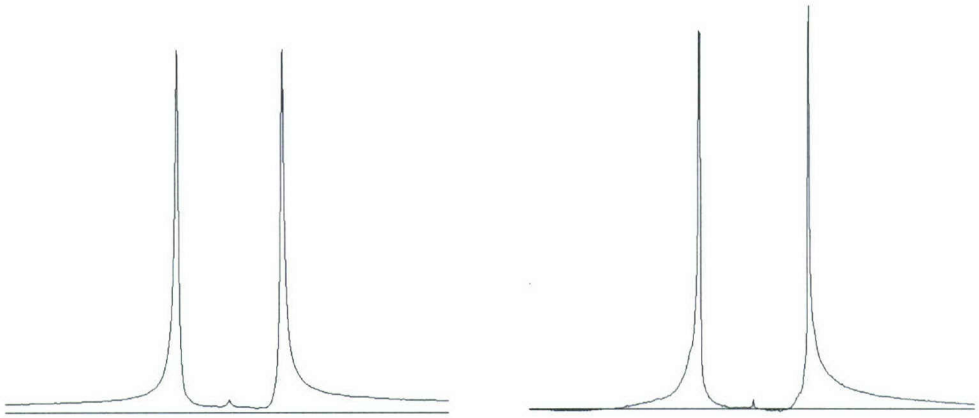
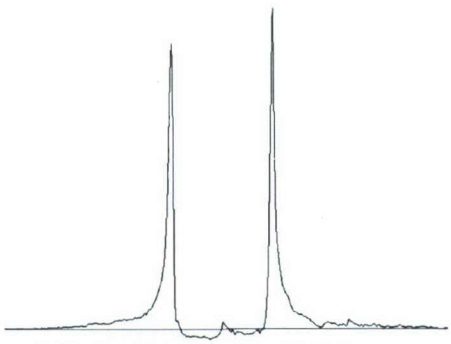


Fig. 11. A four coil array. Helmholtz coil pairs were implemented on 22 mm outer diameter G10 fiberglass tubing and mounted within 38 mm outer diameter RF copper shields. The copper shields which were constructed of electrically isolated strips of copper overlaid on fiberglass tube to form a complete physical barrier. The four coil and shield units were then mounted on two plexi-glass discs of 20.8 cm diameter that would slide within the 21 cm bore of the 5 Tesla magnet and maintain consistent coil placement from experiment to experiment. Holes within the plexi-glass discs allowed for sample placement within the coils and connections of the coils to the tuning boards. Each coil pair was tuned to 212.5 and 53.4 MHz, the Larmor resonant frequencies of H^1 and C^{13}, respectively.



1.5 Tesla, Single Coil, Centered 5 Tesla, Single Coil, Centered



5 Tesla, Coil 1 of the Coil Array,
Offset 4.5 cm from the Magnet's Central Axis

Fig. 12. ^{1}H spectra for the magnets and coil positions used for this work. The upper left shows the spectra in the 1.5 Tesla magnet with the coil centered. The upper right shows the results in the 5 Tesla magnet, again with the coil centered. Finally, the bottom image shows the results for one of the coils in coil array, where the coil itself is off-center from the magnet's central axis, nominally the region of maximum B_0 uniformity. The full-width, half maximum (FWHM) line widths for each of these is approximately 10, 5 and 7 Hz, respectively.

here showed that the incomplete elimination of off-axis elements for pseudo-pure start state appeared to be the dominant error term. This approach relies on gradients to eliminate these off-axis elements. However, eddy currents caused by gradients can interfere with the temporal field homogeneity, distorting the results of the calculation sequence beyond use. For this reason our pulse power of the existing gradient coil magnet insert was limited. The data presented here shows a compromise between gradient coil power and distortions caused by field inhomogeneities due to eddy currents. One approach to address this issue is to use individual gradient

coils around each processor coil for future hardware implementations. The increased distance between the gradient coils and the magnet walls would allow for much stronger gradient pulses without the penalties of eddy current induced field distortion and would have a significant impact on the results.

The incomplete elimination of the off-axis terms is also significant in that this had a multiplicative effect on other error terms. One example is finding that the resultant errors for flip angle inaccuracies are double the expected values when a high degree of off-axis terms remained. This amplification effect is also seen to a lesser degree for evolution errors. A key indicator that off-axis terms are present is the non-symmetric growth of error in each channel. This behavior is generally not seen with the other error terms studied, with the possible exception of random noise, which is at least a magnitude smaller in effect.

Overall signal loss, or decoherence plays a key role in the appearance of the data. Even small overall signal loss (one percent or less) from time of density encoding to final read resulted in the appearance of an asymptotic limit to the density values over several iterations. This signature shape was not seen from other error terms studied.

Finally, the system was equally sensitive to random system errors in both the flip angles and evolution terms. Small magnitudes of one half of one percent random error for flip and evolution terms in the model matched the experimental results well.

5. CONCLUSION

The experimental results show reasonably good agreement with theoretical prediction as well as with previous work using an imaging approach for a type II quantum processor. Simulation of the experimental implementation showed excellent agreement with experimental results. Furthermore, the simulation pointed to a high degree of remaining off-axis magnetization terms from the pseudo-pure state preparation as the main contributor of experimental data deviation from the analytical solution. This indicates that an improved implementation the pseudo-pure state, through the improved implementation of high power, low field-distorting gradients, would greatly improved the quality of experimental data.

Further work is in the development an implementation of multiple-coil type II quantum computer system. This system will use the existing custom software already developed along with the custom two full channel spectrometer but will run on a 5 Tesla magnet and employ smaller coils along with a custom constructed hardware switch to control the operation of the

coils. This switch will allow full control of the coils by the software and allow optimization of coil use based on sample fidelity and experimental needs. The additional coils and switch give the advantage of additional computing power with relatively small expensive as compared to adding additional spectrometer channels as well as additional coils. The improved filling factor of the planned smaller coil design, the increased system field strength and the planned modification to the collision operator[26] should result in improved system fidelity and overall system accuracy.

A four coil insert constructed for the multi-coil work is shown in Fig. 11. This insert has a maximum coil off-set of 45 mm from magnet center. This configuration yields acceptable line widths, even off-axis, as shown in Fig. 12.

ACKNOWLEDGMENTS

We would like to acknowledge his indebtedness to the Air Force Office of Scientific Research (AFOSR) and their support of the Quantum Computation for Physical Modeling theme. Long-term support was provided through the Computational Mathematics Program of the AFOSR for algorithm development research beginning in 1992 under the Novel Strategies for Parallel Computing Initiative at the Air Force Research Laboratory.

REFERENCES

1. Jeffrey Yepez, *Int. J. Modern Phys. C* **12**(9), 1273 (2001).
2. Jeffrey Yepez, *Phys. Rev. A* (2006). (To appear).
3. Jeffrey Yepez, *Int. J. Modern Phy. C* **12**(9), 1285 (2001).
4. G. P. Berman, A. A. Ezhov, D. I. Kamenev, and J. Yepez, *Phy. Rev. A* **66**(012310), 8 (2002).
5. Marco A. Pravia, Zhiying Chen, Jeffrey Yepez, and David G. Cory, *Comp. Phys. Commun.* (2001).
6. Jeffrey Yepez, *J Stat. Phys.* **107**(1), 203 (2002).
7. Zhiying Chen, Jeffrey Yepez, and David G. Cory, *Phys. Rev. A* (2005) arXiv:quant-ph/0410198. (submitted).
8. J. H. Noggle and R. E. Schirmer, *The Nuclear Overhauser Effect, Chemical Applications*, (Academic Press, NY, 1971).
9. A. G. Webb, *Prog. Nucl. Magn Reson. Spectrosc.* **31**, 1 (1997).
10. A. Haase, F. Odoj, M. Von Klienlin, I. Warnking, F. Fidler, A. Weisser, M. Nittka, E. Rommel, T. Lanz, B. Kalusche, and M. Griswold, *Concepts Magn. Reson.* **12**(6), 361 (2000).
11. D. I. Hoult and P. C. Lauterbur, *J. Magn. Reson.* **34**(2), 425 (1979).
12. T. L. Peck, R. L. Magin, and P. C. Lauterbur, *J. Magn. Reson. B* **108**, 114 (1995).

13. Jeffrey Yepez, *Quant. Inform. Proc.* DOI: 10.1007/s11128-005-0009-7.
14. Jeffrey Yepez, *Quant. Inform. Proc.* DOI: 10.1007/s11128-005-0008-8.
15. Jeffrey Yepez, *Int. J. Modern Phy. C* **9**(8), 1587 (July 1998).
16. Marco A. Pravia, PhD. thesis, Massachusetts Institute of Technology(2002).
17. Marco A. Pravia, Zhiying Chen, and David G. Cory, arXiv:quant-ph (Mar 2003).
18. Malcolm H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (Wiley, 2001).
19. F. Bloch, *Phys. Rev.* **70**(7 and 8), 469 (Oct 1946).
20. David G. Cory, Mark D. Price, and Timothy F. Havel. *Phys D* **120**, 82 (1998).
21. Isaac L. Chuang, Lieven M. K. Vandersypen, Xinlan Zhou, Debbie E. Leung, and Seth Lloyd, *Nature* **393**, 143(4) (May 1998).
22. Mark D. Price, Evan M. Forunato, Marco A. Pravia, Craig Breen, Swami Kumaresean, Gabriel Roseberg, and David G. Cory, *Concepts Magn. Reson.* **13**(3), 151 (2001).
23. Marco A. Pravia, Evan Fortunato, Yaahov Weinstein, Mark D. Price, Grum Teklemariam, Richard J. Nelson, Yehuda Sharf, Shyamal Somaroo, C. H. Tseng, Timothy F. Havel, and David G. Cory, *Concepts Magn. Reson.* **11**, 225 (1999).
24. Neil A. Gershenfeld and Isaac L. Chuang, *Science* **275**, 350 (Jan 1997).
25. M. D. Price, S. S. Somaroo, C. H. Tseng, J. C. Gore, A. F. Fahmy, T. F. Havel, and D. G. Cory, *J. Magn. Reson.* **140**(2), 371 (1999).
26. Zhiying Chen, Jeffrey Yepez, and David G. Cory, arXiv:quant-ph, 0410198 (April 2004).

Quantum Information Processing, Vol. 4, No. 6, December 2005 (© 2006)
DOI: 10.1007/s11128-005-0008-8

Lattice Quantum Algorithm for the Schrödinger Wave Equation in 2+1 Dimensions with a Demonstration by Modeling Soliton Instabilities

Jeffrey Yepez,[1,4] George Vahala,[2] and Linda Vahala[3]

Received July 8, 2005; accepted December 5, 2005; published online February 24, 2006

A lattice-based quantum algorithm is presented to model the non-linear Schrödinger-like equations in 2 + 1 dimensions. In this lattice-based model, using only 2 qubits per node, a sequence of unitary collide (qubit–qubit interaction) and stream (qubit translation) operators locally evolve a discrete field of probability amplitudes that in the long-wavelength limit accurately approximates a non-relativistic scalar wave function. The collision operator locally entangles pairs of qubits followed by a streaming operator that spreads the entanglement throughout the two dimensional lattice. The quantum algorithmic scheme employs a non-linear potential that is proportional to the moduli square of the wave function. The model is tested on the transverse modulation instability of a one dimensional soliton wave train, both in its linear and non-linear stages. In the integrable cases where analytical solutions are available, the numerical predictions are in excellent agreement with the theory.

KEY WORDS: Non-linear Schrödinger wave equation; quantum algorithm; soliton dynamics; non-linear quantum mechanical instability; quantum computing; computational physics.

PACS: 03.67.Lx; 05.45.Yv; 02.60.Cb.

1. INTRODUCTION

The non-linear Schrödinger (NLS) equation is one of the most basic equations of non-linear physics. Its salient feature is that it emits soliton solutions by exact integration. Hence, it plays a vital role in weakly non-linear systems

[1] Air Force Research Laboratory, Hanscom Field, Bedford, MA 01731, USA.
[2] Department of Physics, William & Mary, Williamsburg, VA 23187, USA.
[3] College of Engineering & Technology, Old Dominion University, Norfolk, VA 23529, USA.
[4] To whom correspondence should be addressed. E-mail: jeffrey.yepez@hanscom.af.mil, URL:http://qubit.plh.af.mil

with the dispersion relation dependent on the wave amplitude. The NLS equation is pivotal in non-linear optics,[1] plasma physics[2] as well as in ideas for information transfer in optical computers.[3,4] In 1+1 dimensions, both the focusing and defocusing NLS equations are exactly integrable and exhibit soliton solutions. Here, we develop and test a quantum lattice representation of the (focusing) NLS equation in 2+1 dimensions

$$i\partial_t \psi + \partial_{xx}\psi + \partial_{yy}\psi + 2|\psi|^2\psi = 0. \tag{1}$$

building on our previous quantum lattice representations of the Shrödinger wave equation,[5] NLS equation[6] and the vector Manakov system[7] in 1+1 dimensions and the Dirac equation in 3+1 dimensions.[8] In particular, we shall consider the transverse modulational instability of the one-dimensional soliton wave train solution $\psi_\circ(x,t)$ of (1).

In the quantum lattice algorithm for the Schrödinger equation, at each spatial node, the wave function is represented by the interference sum of probability amplitudes of the upper excited state of each qubit. In the quantum algorithm for the NLS equation in 1+1 dimensions, 2 qubits per node are used, and collisional interaction is induced by the unitary $\sqrt{SWAP}$ quantum logic gate. The $\sqrt{SWAP}$ quantum gate has been implemented experimentally using 2 qubits per computational node in a quantum lattice gas model of the diffusion equation.[9,11] This entanglement is then spread throughout the lattice by the unitary streaming operator, which is real. In extending the algorithm to 2+1 dimensions, still only 2 qubits per node are required and the $\sqrt{SWAP}$ still represents local qubit–qubit interactions. If implemented on a platform using future quantum information processing device technology, the quantum algorithm presented here is suited for a type-I quantum computer architecture, as described in this Quantum Computation for Physical Modelling (QCPM) special issue in Section A of,[8] but with local non-linear interactions inherent in the quantum device.

2. QUBIT REPRESENTATION FOR THE NLS WAVE FUNCTION IN 2+1 DIMENSIONS

We discretize the single-particle wave function over a two dimensional square Bravais lattice (the wave function is defined only on a spacetime lattice) where 2 qubits are used at each lattice node to encode the local value of the wave function at that node. Let L denote the number of lattice nodes along an orthogonal direction and let i and j be integer valued spatial indices ranging from 1 up to L. Then, at lattice node (i,j) one defines a position basis ket $|x_{ij}\rangle$. The discretized single-particle wave

function ket $|\psi\rangle$ is modelled by a sum over all possible ways the particle can be located on the lattice sites:

$$|\psi\rangle = \sum_{i,j=0}^{L-1} \gamma^{ij} |x_{ij}\rangle, \tag{2}$$

where the (complex) probability amplitude for each possibility is $\gamma^{ij} \equiv \langle x^{ij} | \psi \rangle$.

The two qubit kets for each lattice node are denoted by $|q_0^{ij}\rangle$ and $|q_1^{ij}\rangle$ with each qubit having the standard two-level representation

$$|q_a^{ij}\rangle = \alpha_a^{ij} |0\rangle + \beta_a^{ij} |1\rangle, \tag{3}$$

with normalization $|\alpha_a^{ij}|^2 + |\beta_a^{ij}|^2 = 1$ for $a = 0, 1$ at spatial site (i, j). In particular, the quantum particle is said to occupy the ath local state at position x_{ij} when $\beta_a^{ij} = 1$, while the ath local state at x_{ij} is empty when $\beta_a^{ij} = 0$. For each position ket there are four basis states in the number representation:

$$\begin{aligned}
&|q_0^{11} q_1^{11}\rangle \cdots \underbrace{|11\rangle}_{x_{ij}} \cdots |q_0^{LL} q_1^{LL}\rangle && \text{doubly occupied at } x_{ij} \\
&|q_0^{11} q_1^{11}\rangle \cdots \underbrace{|10\rangle}_{x_{ij}} \cdots |q_0^{LL} q_1^{LL}\rangle && \text{spin-up at } x_{ij} \\
&|q_0^{11} q_1^{11}\rangle \cdots \underbrace{|01\rangle}_{x_{ij}} \cdots |q_0^{LL} q_1^{LL}\rangle && \text{spin-down at } x_{ij} \\
&|q_0^{11} q_1^{11}\rangle \cdots \underbrace{|00\rangle}_{x_{ij}} \cdots |q_0^{LL} q_1^{LL}\rangle && \text{empty at } x_{ij},
\end{aligned}$$

where we use conventional terminology letting $|q_0^{ij}\rangle$ encode *spin-up* and $|q_1^{ij}\rangle$ encode *spin-down*, say.

In the number representation of the one-particle wave function ket $|\psi\rangle$, we need consider that subset of basis states in which only one amplitude $\beta_a^{ij} = 1$ is non-zero (all other β amplitudes are zero). This subset of basis states is called the *one-particle sector*. There are $(2L)^2$ such states. So in the one-particle sector, there are two ways (interfering possibilities) for a particle to occupy the ijth lattice position

$$\beta_0^{ij} \alpha_1^{ij} |00 \cdots \underbrace{10}_{x_{ij}} \cdots 00\rangle + \alpha_0^{ij} \beta_1^{ij} |00 \cdots \underbrace{01}_{x_{ij}} \cdots 00\rangle. \tag{4}$$

Hence the occupancy probability of the ijth node is determined by first summing up the probability amplitudes of the spin-up and spin-down

basis states in the one-particle sector and then computing the resulting square of this absolute value. Letting $\gamma_{\uparrow}^{ij} \equiv \beta_0^{ij}\alpha_1^{ij}$ and $\gamma_{\downarrow}^{ij} \equiv \alpha_0^{ij}\beta_1^{ij}$, the complex probability amplitude in (2) is set equal to the sum of the two on-site probability amplitudes

$$\gamma_{ij} = \gamma_{\uparrow}^{ij} + \gamma_{\downarrow}^{ij}. \tag{5}$$

3. QUANTUM ALGORITHM TO RECOVER THE NON-LINEAR SCHRÖDINGER EQUATION IN 2+1 DIMENSIONS

To recover a macroscopic scale effective theory that approximates the Schrödinger wave equation in the long wave length limit, our quantum lattice representation of the dynamics uses the unitary $\sqrt{SWAP}$ quantum logic gate as the collision operator that couples the on-site probability amplitudes:

$$\mathscr{C}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-i & 1+i\\ 1+i & 1-i\end{pmatrix}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix}, \tag{6}$$

and eight stream operators which independently shift the $\uparrow$ and $\downarrow$ components of the discretized spinor wave function in the $\pm\hat{x}$ and $\pm\hat{y}$ directions. The stream operator and its transpose (which is its adjoint and inverse) in the $\hat{x}$ direction for the first (spin-up) component are:

$$S_{x\uparrow}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i+1,j}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} \qquad S_{x\uparrow}^{\mathrm{T}}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i-1,j}\\ \gamma_{\downarrow}^{ij}\end{pmatrix}, \tag{7}$$

and the stream operators in the $\hat{x}$ direction for the second (spin-down) component of the discretized spinor wave function:

$$S_{x\downarrow}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i,j}\\ \gamma_{\downarrow}^{i+1,j}\end{pmatrix} \qquad S_{x\downarrow}^{\mathrm{T}}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i,j}\\ \gamma_{\downarrow}^{i-1,j}\end{pmatrix}. \tag{8}$$

Similarly, we define the four stream operators in the $\hat{y}$ direction:

$$S_{y\uparrow}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i,j+1}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} \qquad S_{y\uparrow}^{\mathrm{T}}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{i,j-1}\\ \gamma_{\downarrow}^{ij}\end{pmatrix}. \tag{9}$$

$$S_{y\downarrow}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{i,j+1}\end{pmatrix} \qquad S_{y\downarrow}^{\mathrm{T}}\begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{ij}\end{pmatrix} = \begin{pmatrix}\gamma_{\uparrow}^{ij}\\ \gamma_{\downarrow}^{i,j-1}\end{pmatrix}. \tag{10}$$

We define the fundamental evolution operator for direction $\hat{w}=\hat{x}$ or $\hat{y}$ and spin $\sigma=\uparrow$ or $\downarrow$ as follows:

$$\mathcal{U}_{w\sigma}=S_{w\sigma}\mathcal{C}. \tag{11}$$

Now we define the *interleaved evolution operator* (which would be identity if the stream and collide operators commuted) as follows:

$$\begin{aligned}\mathcal{J}_{w\sigma} &\equiv S_{w\sigma}^{\dagger}\mathcal{U}_{w\sigma}^{\dagger}S_{w\sigma}\mathcal{U}_{w\sigma} && \text{(12a)}\\ &= S_{w\sigma}^{\dagger}\mathcal{C}^{\dagger}S_{w\sigma}\mathcal{C} && \text{(12b)}\\ &= S_{w\sigma}^{\mathrm{T}}\mathcal{C}S_{w\sigma}\mathcal{C} && \text{(12c)}\\ &= S_{-w,\sigma}\mathcal{C}S_{w\sigma}\mathcal{C}, && \text{(12d)}\end{aligned}$$

since the adjoint of the fundamental evolution operator is $\mathcal{U}_{w\sigma}^{\dagger}=\mathcal{C}^{\dagger}S_{w\sigma}^{\dagger}$, the stream operator is real $S_{w\sigma}^{\dagger}=S_{w\sigma}^{\mathrm{T}}$, and the collide operator is self-adjoint $\mathcal{C}^{\dagger}=\mathcal{C}$. Because of the spacetime interpretation of spin,[8] where if spin-up moves along $\hat{w}$ say then spin-down moves along $-\hat{w}$, then the interleaved evolution operator is invariant under the following simultaneous spin flip and spatial inversion:

$$\mathcal{J}_{-w,-\sigma}=\mathcal{J}_{w\sigma}. \tag{13}$$

For example, $\mathcal{J}_{-x,\uparrow}=\mathcal{J}_{x\downarrow}$.

Now, let E denote the local quantum evolution operator that advances the discretized spinor wave function one unit in time. Then the evolution equation is the following:

$$\begin{pmatrix}\gamma_{\uparrow}^{i,j}(t+\Delta t)\\ \gamma_{\downarrow}^{ij}(t+\Delta t)\end{pmatrix}=E\begin{pmatrix}\gamma_{\uparrow}^{i,j}(t)\\ \gamma_{\downarrow}^{ij}(t)\end{pmatrix} \tag{14}$$

The evolution operator can be partitioned in space using an operator splitting method. A third-order accurate quantum algorithm for the local evolution operator has the form

$$E=\mathcal{J}_{-y\downarrow}\mathcal{J}_{y\uparrow}\mathcal{J}_{-x\uparrow}\mathcal{J}_{x\downarrow} \tag{15}$$

where the macroscopic effective field theory for the spinor field $\psi=\begin{pmatrix}\gamma_{\uparrow}\\ \gamma_{\downarrow}\end{pmatrix}$ is

$$\partial_t\psi=\frac{i}{2}\frac{\Delta x^2}{\Delta t}\sigma_x\left(\partial_{xx}+\partial_{yy}\right)\psi+\mathcal{O}(\epsilon^3), \tag{16}$$

where $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and where $\epsilon \sim \Delta x \sim \sqrt{\Delta t}$. So if we trace over the spin components of ψ and form the scalar $\Psi \equiv \gamma_\uparrow + \gamma_\downarrow$, we obtain the following non-relativistic wave equation for a free quantum particle:

$$\partial_t \Psi = i \frac{\hbar}{2m} \left(\partial_{xx} + \partial_{yy} \right) \Psi + \mathcal{O}(\epsilon^3), \tag{17}$$

where the diffusion constant associated with the particle mass is $(\hbar/m = \Delta x^2/\Delta t)$ in lattice units. Quadratic products of the interleaved evolution operator (12) are invariant to order ϵ^3 under the following double spin flip and spatial interchange operation:

$$\mathcal{J}_{w,\sigma} \mathcal{J}_{w',\sigma'} = \mathcal{J}_{w',-\sigma'} \mathcal{J}_{w,-\sigma} + \mathcal{O}(\epsilon^3). \tag{18}$$

For example, $\mathcal{J}_{x\uparrow} \mathcal{J}_{y\downarrow} = \mathcal{J}_{y\uparrow} \mathcal{J}_{x\downarrow}$ and $\mathcal{J}_{x\uparrow} \mathcal{J}_{y\uparrow} = \mathcal{J}_{y\downarrow} \mathcal{J}_{x\downarrow}$. Hence using (18), there are $\binom{4}{2} = 6$ ways to re-order the spatial indices of the evolution operator (15). Then, using (13), for each configuration of the spatial indices there are 16 ways to re-order the spin indices of (15). Hence, there are a total of 96 ways to rewrite the quantum algorithm (15). For example, we can rewrite (15):

$$E \overset{(13)}{=} \mathcal{J}_{y\uparrow} \mathcal{J}_{y\uparrow} \mathcal{J}_{x\downarrow} \mathcal{J}_{x\downarrow} \tag{19a}$$

$$\overset{(18)}{=} \mathcal{J}_{y\uparrow} \mathcal{J}_{x\uparrow} \mathcal{J}_{y\downarrow} \mathcal{J}_{x\downarrow} \tag{19b}$$

$$\overset{(18)}{=} \mathcal{J}_{x\uparrow} \mathcal{J}_{x\uparrow} \mathcal{J}_{y\downarrow} \mathcal{J}_{y\downarrow}. \tag{19c}$$

Every version has the same algorithmic complexity and the error terms are always order ϵ^3. Furthermore, there are versions of the quantum algorithm where the error terms differ only by an overall sign change. We exploit this feature to judiciously cause a cancellation of all ϵ^3 error terms by using twice as many operators. This doubles the algorithmic complexity, but the error is then pushed out to fourth-order. Although the algorithmic complexity increases by a factor of 2, the numerical accuracy of the algorithm increases by a factor of 4 because of the diffusive ordering of the space and time fluctuations (because (17) is parabolic). Therefore, it is advantageous to employ this numerical schema.

As a case in point, the error terms in (19a) and in (19c) differ only by an overall sign. Hence, choosing our evolution operator to be

$$E = \mathcal{J}^2_{x\uparrow} \mathcal{J}^2_{y\downarrow} \mathcal{J}^2_{y\uparrow} \mathcal{J}^2_{x\downarrow}, \tag{20}$$

we recover the following macroscopic effective field theory for the spinor field

$$\partial_t \psi = i \frac{\Delta x^2}{\Delta t} \sigma_x \left(\partial_{xx} + \partial_{yy} \right) \psi + \mathcal{O}(\epsilon^4). \tag{21}$$

Again, tracing over the spin degrees of freedom, we obtain the Schrödinger wave equation as our effective field theory, but now with diffusion constant

$$\frac{\hbar}{m} = 2 \frac{\Delta x^2}{\Delta t}. \tag{22}$$

We add a potential V by rotating the overall phase of the spinor field following each application of (20)

$$\psi(t + \Delta t) = E e^{-i \Delta t V/\hbar} \psi(t) \tag{23}$$

$$\overset{(22)}{=} E e^{-i \Delta x^2 (2m/\hbar^2) V} \psi(t). \tag{24}$$

The resulting equation of motion is

$$\partial_t \psi = i \frac{\Delta x^2}{\Delta t} \sigma_x \left(\partial_{xx} + \partial_{yy} \right) \psi - \frac{i}{\hbar} V \psi + \mathcal{O}(\epsilon^4), \tag{25}$$

or in terms of the scalar wave function

$$i \hbar \partial_t \Psi = -\hbar \frac{\Delta x^2}{\Delta t} \left(\partial_{xx} + \partial_{yy} \right) \Psi + V \Psi + \mathcal{O}(\epsilon^4). \tag{26}$$

The addition of the potential does not introduce any greater error nor diminishes the numerical accuracy of the scheme. Using (13), there are 2^8 ways to rewrite (20). Furthermore, (20) must be invariant under an interchange of the spatial labels x and y and the spin labels $\uparrow$ and $\downarrow$. Hence, there are at least 256 ways of writing a quantum algorithm that is fourth-order accurate.

4. TRANSVERSE INSTABILITY

For convenience, we briefly review some properties of a 1D soliton wave train solution of (1).[1] A planar 1D bright soliton solution of (1) is

$$\psi_0(x, t) = \phi(x - x_0 - 2vt; \beta) e^{i(vx - v^2 t + \beta t + \theta)}, \tag{27}$$

where the standard soliton shape is given by

$$\phi(x;\beta) = \beta^{1/2}\mathrm{sech}(\beta^{1/2}x) \tag{28}$$

The location of the soliton wave train is x_0, $2v$ is the (transverse) soliton wave train speed, β controls its amplitude and θ its phase. For a linear stability analysis of this 1D soliton wave train, one considers perturbations of the form

$$\delta\psi(x, y, t) = [u(x) + i\,w(x)]e^{i\beta t + \Gamma t + ipy}, \tag{29}$$

where p is the transverse perturbation wavenumber in the y-direction, and $\Gamma(p)$ is the linear growth rate of the perturbation. An analytic solution to this linear perturbation problem does not exist, and one must resort to asymptotic theory—either about the long wavelength limit ($p=0$) or about the maximum growth rate wavenumber p_c where $\Gamma(p_c)=0$. Here, we shall consider the linear instability in the long wavelength limit, $p \ll 1$. In this limit, it can be shown[1] that the resulting eigenvalue problem has

$$u(x)=0 \qquad w(x)=\phi(x). \tag{30}$$

In the long wavelength limit, the transverse modulation will break up the 1D soliton wave train into N filaments, where

$$N = \frac{pL_y}{2\pi}. \tag{31}$$

5. NUMERICAL PREDICTIONS FOR NLS EQUATION IN 2+1 DIMENSIONS

We apply our quantum lattice algorithm to the solution of the NLS equation in 2+1 dimensions, using a 1024×1024 spatial grid with a soliton wave train speed $v=0.05$ and amplitude $\beta^{1/2}=0.085$ for three cases presented below: an unperturbed soliton wave train, a solition wave train with an additional transverse modulation, and the interaction of two perpendicularly directed solition wave trains.

To test whether (classical) floating-point roundoff will trigger the transverse instability in the quantum lattice algorithm we the propagation of the planar 1D bright soliton solution in the NLS equation in 2+1 dimensions. We find even after 10,000 lattice time steps there is no triggering of the transverse modulation instability, and the 1D soliton train

propagates undistorted through the lattice, see Fig. 1. The quantum algorithm is numerically stable. This is verified in the simulations for $t = 0\Delta t$, $t = 5\,\mathrm{K}\Delta t$, $t = 10\,\mathrm{K}\Delta t$. The simulation error in soliton speed after 10K iterations is only 0.5% (about 5 lattice grid points on the 1024 grid).

A transverse perturbation of the form of (29) with wavenumber $p = 2\pi N/L_y$, $N = 8$ with an amplitude that is a factor of 10^{-7} below the initial wave train amplitude. The asymptotic linear stability predicts that the 1D soliton train should break up into 8 filamentary structures, an unstable phase, as shown in Fig. 2. This is indeed found in our simulations, with the filamentary structures becoming so isolated and peaked that the steep gradient of the wave function cannot be resolved on the grid after about $t = 2100\Delta t$ iterations, even on using a 1024×1024 lattice.

Finally, we consider the interaction of two perpendicularly directed soliton wave trains. No initial perturbations are needed due to the "overlap" region of the two wave trains. A very rapid instability is immediately triggered at the intersection region of the two wave trains, and its localization and peak are so rapid that the gradient of the wave function cannot be further resolved on the grid after just 35 time steps. The instability occurs only at the point of intersection of the two soliton trains and the unaffected regions of both soliton trains propagate in normal fashion. This expected behaviour is observed in the simulation shown in Fig. 3.

6. CONCLUSION

Presented was a lattice-based quantum algorithmic method to numerically model time-dependent solutions of the Schrödinger wave equation in an arbitrary number of spatial dimensions using a fourth-order accurate operator splitting method. Here, we tested the method in 2+1 dimensions using a quantum system with a non-linear potential. Generalization to 3+1 dimensions is straightforward. This was the first numerical test of the quantum lattice gas algorithm in multiple spatial dimensions—all previous simulation results that have appeared in the literature have been for 1+1 dimensional cases. Furthermore, we probed to determine if the quantum algorithm was cable of accurately modelling the expected physical behaviour of the non-linear quantum system by triggering the onset of strong and rapid non-linear instabilities in solitary wave trains. This was a stringent test of the method.

In all the cases, the quantum algorithm performed excellently with the numerical results in perfect agreement with the theoretical predictions. Ultimately, in tracking the late time developed of the growth of the instabilities, we were limited by our fixed grid resolution. To follow

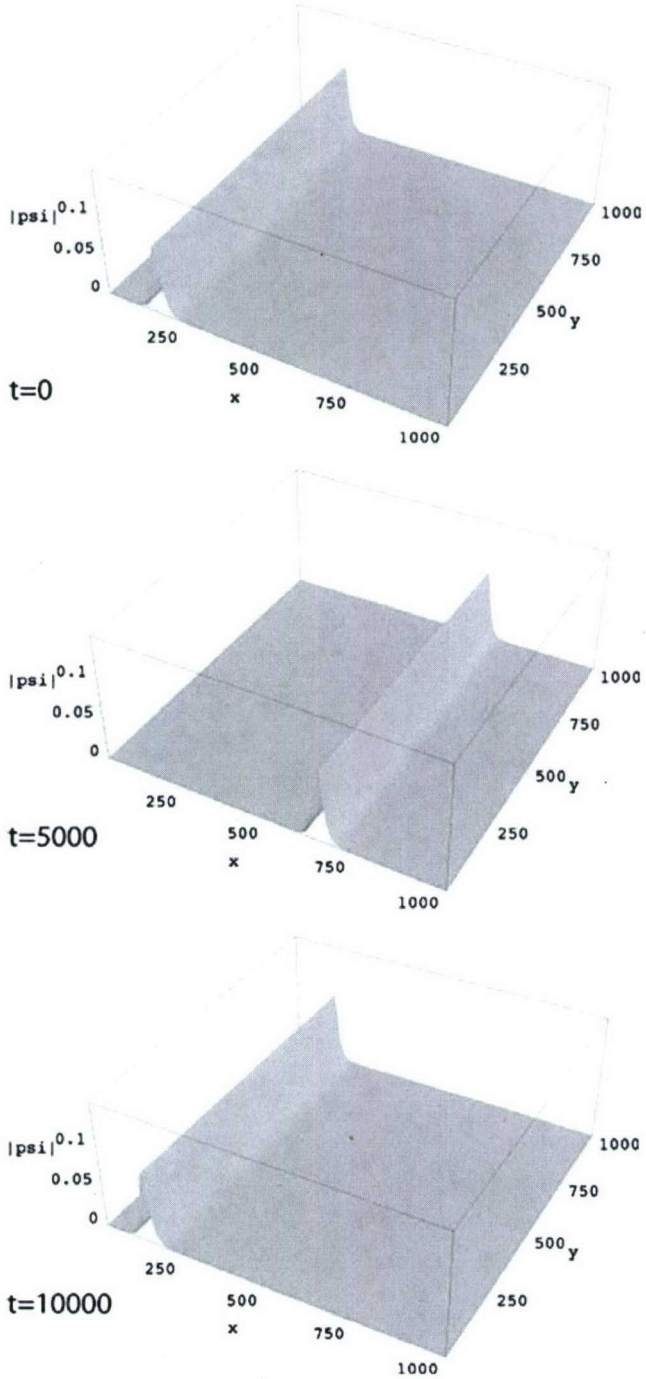


Fig. 1. Evolution of a 1D soliton wave train for the NLS equation in 2+1 dimensions on a 1024^2 grid with periodic boundary conditions. No transverse modulation instabilities are triggered, even after 10,000 iterations. By $t = 10,1000\Delta t$ (bottom), the wave train has wrapped around the grid.

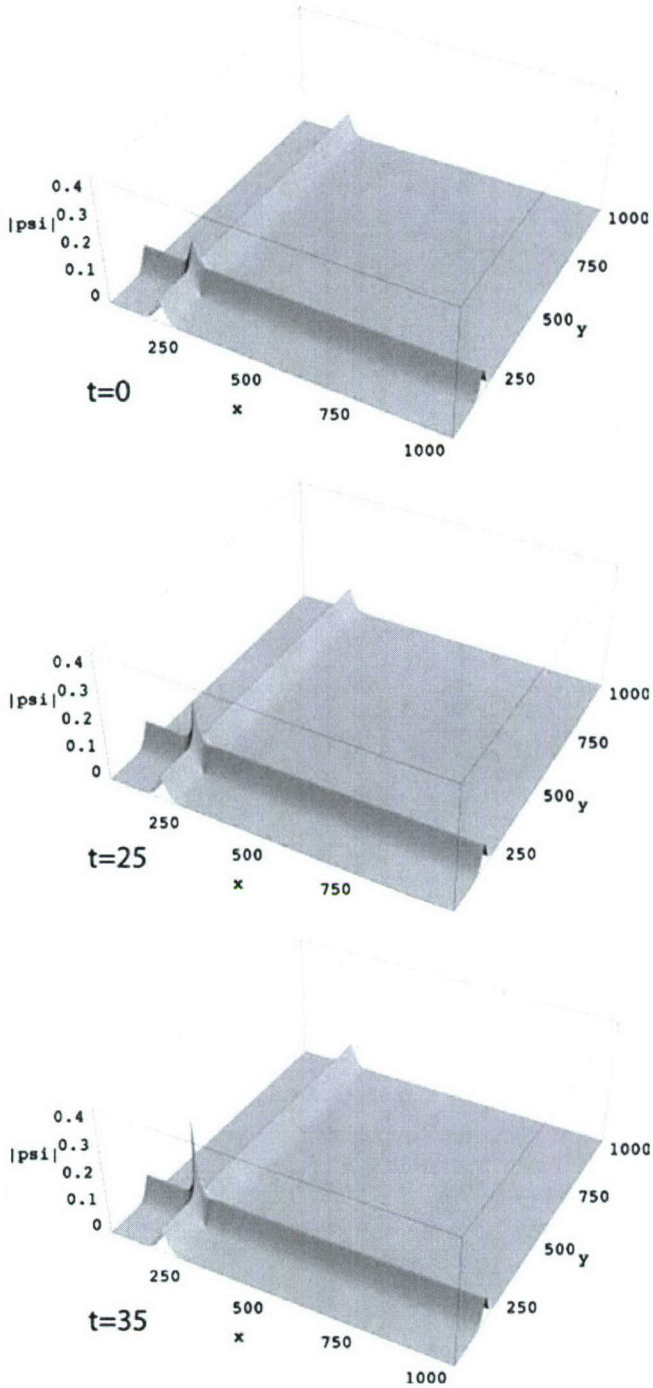


Fig. 2. Evolution of a 1D soliton wave train for the NLS equation in 2+1 dimensions on a 1024^2 grid (only half the grid is shown) with a transverse perturbation with amplitude 10^{-7} lower than the initial peak amplitude of the soliton wave train. A transverse modulation instability is triggered, clearly observable after $t = 2000\Delta t$ time steps.

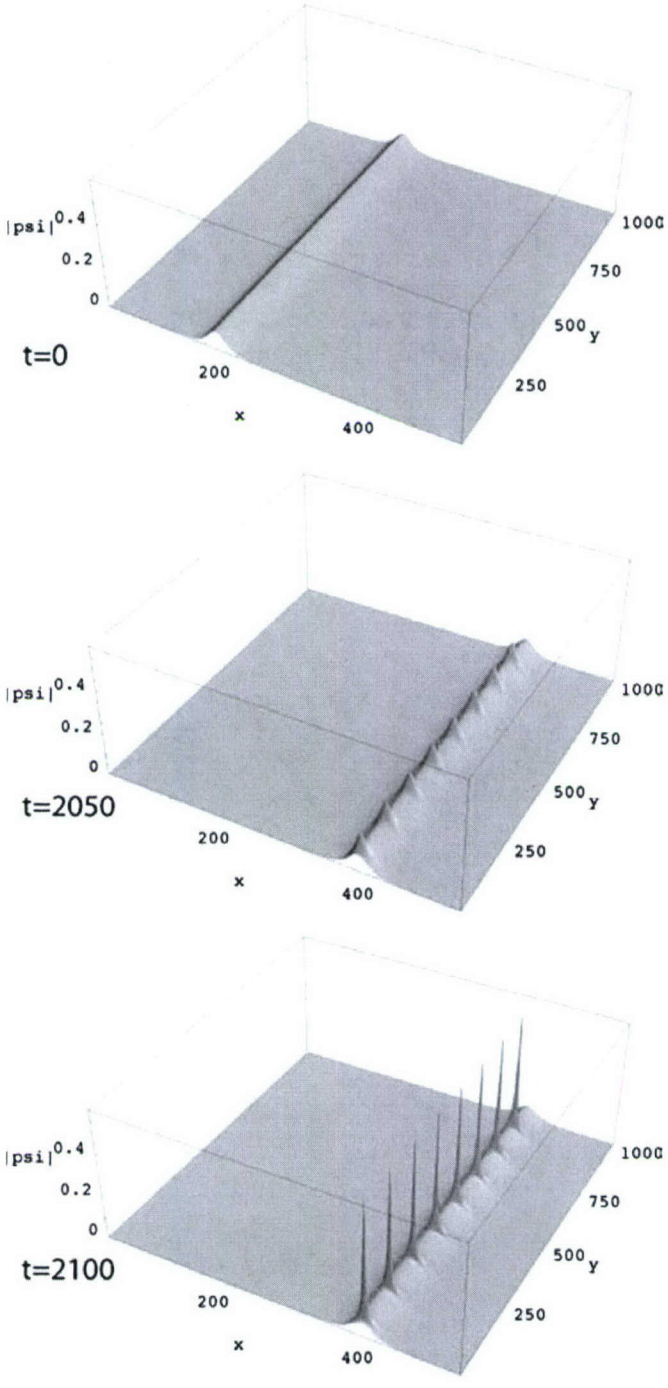


Fig. 3. Evolution of a two orthogonally directed 1D soliton wave trains for the NLS equation in 2+1 dimensions on a 1024^2 grid. An rapid instability is immediately triggered, creating a rising peak at the intersection point of the solitons that reached the grid resolution after $t = 35\Delta t$ time steps.

the quantum evolution for significantly longer periods of time following the onset of a non-linear instability, one could introduce adaptive mesh refinements into our quantum algorithmic scheme, and this will be left for future work.

ACKNOWLEDGMENT

The authors wish to express their indebtedness for the support provided through the Computational Mathematics Program of the Air Force Office of Scientific Research for algorithm development research.

REFERENCES

1. Y. S. Kivshar and G. P. Agrawal, *Optical Solitons* (Academic Press, New York, 2003).
2. E. Infeld and G. Rowlands, *Nonlinear Waves, Solitons, and Chaos*, 2nd ed. (Cambridge University Press, Cambridge, 2000).
3. M. H. Jakubowski, K. Steiglitz, and R. Squier, *Phys. Rev. E* **58**, 6752 (1998).
4. M. H. Jakubowski, K. Steiglitz, and R. Squier, *Multiple-Valued Logic* **6**, 439 (2001).
5. Jeffrey Yepez, and Bruce Boghosian, *Comput. Phys. Commun.* **146**(3), 280–294 (2002).
6. George Vahala, Linda Vahala, and Jeffrey Yepez, *Phys. Lett. A* **310**, 187–196 (2003).
7. George Vahala, Linda Vahala, and Jeffrey Yepez, *Philosop. Trans. R. Soc.* **362**(1821), 1677–1690 (2004).
8. Jeffrey Yepez, *Quantum Information Processing* (6), (2005) (in press).
9. Marco Pravia, Zhiying Chen, Jeffrey Yepez, and David G. Cory, *Comput. Phys. Commun.* **146**(3), 339–344 (2002).
10. Marco Pravia, Zhiying Chen, Jeffrey Yepez, and David G. Cory, *Quantum Information Processing* **2**, 1–19 (2003).
11. Lisa C. Siskind, Bruce E. Hammer, Nelson L. Christensen, and Jeffrey Yepez, *Quantum Information Processing* **4**(6) (2005). Appearing in this special issue of QIP.

Quantum Information Processing, Vol. 4, No. 6, December 2005 (© 2006)
DOI: 10.1007/s11128-005-0009-7

Relativistic Path Integral as a Lattice-based Quantum Algorithm

Jeffrey Yepez[1,2]

Received August 22, 2005; accepted December 13, 2005; published online February 16, 2006

We demonstrate the equivalence of two representations of many-body relativistic quantum mechanics: the quantum lattice-gas method and the path integral method. The former serves as an efficient lattice-based quantum algorithm to simulate the space-time dynamics of a system of Dirac particles.

KEY WORDS: Dirac equation; quantum computing; quantum lattice gas; many-body relativistic quantum mechanics; path integrals.

PACS: 03.67.Lx; 03.65.Pm; 04.25.Dm; 05.30.Fk.

1. INTRODUCTION

Finding a simple rule to represent the spacetime quantum mechanical dynamics of a system of Dirac particles in 1+1 dimensions as a discrete path integral, or more accurately as a path summation, is known as the *Feynman chessboard problem.*(1) In Feynman's notes we see that he first solved this problem in 1946.(2) A proof by Jacobson and Schulman of Feynman's solution to this chessboard problem relies on a deep isomorphism between the discrete path integral and the partition function in statistical mechanics of an Ising spin system with nearest-neighbor spin–spin interaction.(3) The 1+1 dimensional chessboard is a square spacetime lattice with

[1] This discrete path integral formalism, included in the beginning of this paper, was presented on August 20, 2004 as an invited talk entitled "Lattice-based quantum algorithms for computational phsyics" at the 13th International Conference on the Discrete Simulation of Fluid Dynamics, hosted by Tufts University in Cambridge, Massachusetts. The quantum algorithm for the Dirac system in 3+1 dimensions, included at the end of this paper, was presented on May 9, 2002 at the Quantum Computation for Physical Modeling Workshop 2002, hosted by the Air Force Research Laboratory in Edgartown, Massachusetts.

[2] Air Force Research Laboratory, 29 Randolph Road, Hanscom Field, Massachusetts 01731. E-mail: Jeffrey.Yepez@hanscom.af.mil; URL: http://qubit.plh.af.mil

grid sizes Δz and Δt. Feynman's solution is the following: the probability amplitude for a free massive Dirac particle to go from one lattice site to another is represented by summing over all the possible zigzag pathways a particle may travel with velocity $\pm\Delta z/\Delta t$, hopping $\pm\Delta z$ from lattice site to lattice site, continuing forward or reversing direction at each time step Δt. The probability amplitude a particular path contributes to the kernel depends on the number of reversals, or bends, where each bend contributes a multiplicative factor of $i\frac{mc^2\Delta t}{\hbar}$ to the overall probability amplitude associated with the path and where m is the particle mass.

Employing just this rule, the relativistic quantum mechanical evolution is correctly emulated. The Dirac Hamiltonian in 1+1 dimensions is recovered in the continuum limit as the grid resolution becomes infinite: $\Delta z \to 0$ and $\Delta t \to 0$ where the limit is taken such that the ratio $\Delta z/\Delta t$ remains constant, which is interpreted as the speed of light c. The Dirac equation of motion in 1+1 dimensions for a 2-spinor field emerges as the effective field theory in the long wavelength limit. A solution to the chessboard problem is given in §II, employing the isomorphism between the path summation and the partition function, but also employing the *stream* and *collide* paradigm used in quantum lattice gases to further simplify matters.

Feynman attempted to find a simple solution to the chessboard problem in 3+1 dimensions, but had no success. Jacobson presented a solution, but his model is unsettling, if not unphysical, because the Dirac particle locally moves faster than the speed of light.[4] Furthermore, Jacobson's solution is complicated because it is implemented on a kind of random spacetime lattice and, hence, is neither useful for numerical simulation purposes nor quantum computation. Here we consider a simple solution for 3+1 dimensions where the Dirac particle locally moves at the speed of light and where the solution is directly suited for numerical simulation and quantum computation.

The full problem, checkered in four dimensions instead of on the plane, is the following: Using a spatial body-centered cubic Bravais lattice, show that the probability amplitude for a free massive Dirac particle to go from one lattice site to another, by moving independently and simultaneously along the orthogonal cubic lattice directions with velocity $\pm\Delta x/\Delta t, \pm\Delta y/\Delta t$, and $\pm\Delta z/\Delta t$, is equal to the sum of all possible pathways between those sites, where the probability amplitude a particular path contributes to the kernel depends on the number of reversals of motion, or bends, counted by projecting along the orthogonal $\hat{x}$, $\hat{y}$, and $\hat{z}$ axes, where each bend contributes a multiplicative factor of $i\frac{mc^2\Delta t}{D\hbar}$ to the overall probability amplitude associated with the path, where m is the particle mass, and $D=3$ is the number of spatial dimension.

Presented in §III is a solution that is a straightforward generalization of the 1+1 dimensional solution given in §II. The key is to designate the quantum mechanical velocity vector of the Dirac particle in terms of the spin components $s_x = \pm 1, s_y = \pm 1$, and $s_z = \pm 1$ as follows:

$$\frac{c}{\sqrt{3}}(s_x \hat{x} + s_y \hat{y} + s_z \hat{z}). \tag{1}$$

Locally there are only eight possible spin combinations, hence the choice of the body-centered cubic lattice. All the spacetime translational degrees of freedom are specified in terms of the spin variables. The relativistically invariant Dirac system is recovered in the continuum limit as the grid resolution becomes infinite: $\Delta r \to 0$ and $\Delta t \to 0$ where $\Delta r = \Delta x = \Delta y = \Delta z$ and where the limit is taken such that the ratio $\Delta r / \Delta t$ remains constant, which is interpreted as the speed of light c. The Dirac Hamiltonian generating the unitary evolution for the 4-spinor field of a relativistic particle emerges as the effective field theory in the long wavelength.

1.1. Application to Quantum Computing

The basic approach is to use an externally controlled array of qubits to encode the quantum wave function of the modeled quantum system of Dirac particles and to use an engineered local Hamiltonian to emulate the Dirac Hamiltonian so as to approximate the subsequent time-dependent behavior of the wave function on a spacetime lattice. Two qubits are used per lattice node to model the Dirac equation in 1+1 and 2+1 dimensions and four qubits are used per lattice node to model the Dirac equation in 3+1 dimensions. The 3 dimensional spatial lattice is a body-centered cubic (bcc) lattice. The 3+1 dimensional spacetime lattice is a hyper-bcc lattice.

The final result of the computation is obtained by measuring all Q qubits. To recover the moduli squared of the modeled wave function, an ensemble measurement or repeated measurement is required. The frequency of occurrence is associated with the moduli squared of the wave function. Upon measurement, the probability of finding a "particle" at time t at a lattice node at location $\vec{x}$ is equals the sum of the moduli squared of the excited-energy eigenstate probability amplitudes of the qubits at that quantum node. This probability is called an *occupation probability*.

The local Hamiltonian that emulates the Dirac Hamiltonian has a kinetic energy part related to the motion of qubits and an on-site qubit–qubit interaction part. Both generate the time-dependent dynamics of a discrete amplitude field. The local Hamiltonian is spatially homogeneously

applied to the lattice-based amplitude field independently on all the lattice nodes. Each part of the Hamiltonian generates unitary evolution operators:

1. The on-site interaction, represented by a unitary operator $\mathscr{C}$ called the *collision operator*, emulates particle-particle "collisions."
2. Site-to-site exchange, represented by a unitary operator $\mathscr{S}$ called the *stream operator*, emulates local "translation" of particles between neighboring lattice sites.

The quantum mechanical local evolution equation is:

$$|\psi(\vec{x}, t+\Delta t)\rangle = \mathscr{C}\mathscr{S}|\psi(\vec{x}, t)\rangle, \tag{2}$$

where $|\psi\rangle$ is local value of the probability amplitude field and where $\mathscr{S}$ causes quantum mechanical streaming. Equation (2) is called a *quantum lattice gas equation of motion* and is derived in §II D for the 1+1 dimensional case and in §III D for the 3+1 dimensional case.

In the one particle case, (2) can rewritten as follows:

$$|\psi(\vec{x}, t+\Delta t)\rangle = \mathscr{C} \sum_{\vec{c} \in \text{bcc}} M_S |\psi(\vec{x}-\vec{c}, t)\rangle, \tag{3}$$

where M_S is a matrix that represents $\mathscr{S}$ on a local stencil of the body-centered cubic (bcc) lattice and the sum is over the eight diagonal neighbors. Then with $\mathscr{C}$ represented by a unitary matrix, (3) becomes a finite-difference representation of the Dirac equation.

The first-order numerically convergent version of the quantum algorithm can be expressed as a quantum lattice-gas evolution Eq. (2). In 3+1 dimensions, $\mathscr{S}$ moves each quantum particle located at the lattice site $\vec{x}$ to 8 neighboring sites on the body-centered cubic lattice, which on R.H.S. of (3) is notated $\vec{x}-\vec{c}$.[3] $\mathscr{S}$ is quantum mechanical because a particle moves to multiple locations at each time unit Δt.

In §IV, the unitary matrix representation of the stream and collide quantum mechanical operators are given to specify quantum algorithms for numerically modeling a system of relativistic Dirac particles. The quantum algorithms are distinguished by whether they may handle one or many particles and by their numerical convergence properties. Several quantum algorithms are presented in §IV: one particle algorithms in §IV A in the form of (3) and many particle algorithms in §IV B in the form of (2) in

[3] Say the net motion is along $\hat{z}$-axis. Then a positive energy particle starting at $(0, 0, 0)$ moves to the 4 sites $(\pm\Delta x, \pm\Delta y, \Delta z)$ whereas a negative energy particle moves to $(\pm\Delta x, \pm\Delta y, -\Delta z)$.

the second quantized representation using ladder operators. The structure of $\mathscr{S}$ depends on the number of spatial dimensions of the modeled quantum mechanical system and it can be altered to improve the numerical convergence properties of the quantum algorithm.

The quantum mechanical stream operator is rewritten in terms of a classical stream operator, denoted $S = S_x S_y S_z$ (not to be confused with the action). Each product component operator of S moves a particle along an orthogonal cubic lattice direction. The following identity relating the quantum stream operator to the classical stream operator is proved in §IV A.2:

$$\mathscr{S} = e^{i\frac{\pi}{4}\sigma_y^{(2)}} S_x e^{-i\frac{\pi}{4}(\sigma_y^{(2)}+\sigma_x^{(2)})} S_y e^{i\frac{\pi}{4}\sigma_x^{(2)}} S_z. \tag{4}$$

As a starting point, in §IV A.2, the simplest quantum lattice-gas algorithm on a bcc lattice is treated. This quantum algorithm is a direct transcription of the path summation in §III. However, this algorithm suffers two difficiencies: the checkboard problem of non-interacting sublattices and only first-order numerical convergence. Therefore, next in §IV A.3, an improved version of the quantum algorithm that remedies these deficiencies is presented.

The simulation of many Dirac particles is treated in §IV B. This quantum algorithm is naturally suited to implementation on a quantum computer using conservative quantum gates. If implemented on this hypothethical quantum computer, it efficiently predicts time-dependendent solutions of the four spinor Dirac field associated with the many-body relativistic system in 3+1 dimensions.

The quantum algorithm can handle the many-body Dirac system in a second-quantized representation without any additional computational overhead. Therefore, this is an efficient quantum algorithm for the many-body system of Dirac particles in 3+1 dimensions. Its computational complexity is dominantly linear in the number of qubits used to spatially resolve the 4-spinor wave function. With Q-number of qubits, the number of quantum gate operations needed to model a one-body quantum system is dominantly linear $\sim \alpha Q$, where α is a fixed constant. This is discussed in §IV B.1. The efficiency of the quantum computation derives from the fact that exactly the same number of quantum gate operations are needed to model a many-body quantum system, for up to as many particles as qubits. This kind of exponential speedup has been previously shown to occur in quantum lattice-gas algorithms for modeling a many-body system of non-relativistic particles.[5]

2. 1+1 DIMENSIONS

2.1. Feynman Path Integral Representation of Non-relativistic Quantum Mechanics

The probability amplitude that a quantum particle at position z_a at time t_a will transfer to a new position z_b and time t_b is given by the following path integral:

$$K(z_a t_a; z_b t_b) = \int \mathcal{D}[z(t)] e^{i \frac{S[z(t)]}{\hbar}}, \tag{5}$$

where $\int \mathcal{D}[z(t)]$ denotes integration over all trajectories $z(t)$ for which $z(t_a) = z_a$ and $z(t_b) = z_b$, and where the increase of the action

$$S = \int_{t_a}^{t_b} dt\, L[\dot{z}(t), z(t)], \tag{6}$$

along a trajectory $z(t)$ is determined using the classical Lagrangian L.

2.2. Path Summation Representation of Relativistic Quantum Mechanics

Feynman[1] found a discrete representation of the Feynman path integral principle on a square spacetime lattice to compute (5) for a relativistic quantum particle:

$$K_{\alpha\beta}(z_a t_a; z_b t_b) = \lim_{N \to \infty} \sum_{R \geq 0} \Phi_{\alpha\beta}(R) \left(i \Delta t \frac{mc^2}{\hbar} \right)^R, \tag{7}$$

where $\Delta t \equiv \frac{t_b - t_a}{N}$, where α and β are the $\pm$ components of the spinor amplitude field (spin-up or spin-down), $\Phi_{\alpha\beta}(R)$ is the number of paths with N steps and R bends, where the length of each step is $\Delta z \equiv \frac{z_b - z_a}{M} \equiv c\, \Delta t$, where c is the speed of light, and where m is the mass of the quantum particle. An example relativistic trajectory with 4 bends along $\hat{z}$ is depicted in Fig. 1.

The allowable region of the lattice is bounded by the intersection of two light cones, with boundaries $z = \pm c(t - t_a) + z_a$ and $z = \pm c(t - t_b) + z_b$ for $t_a \leq t \leq t_b$ and $z_a \leq z \leq z_b$. One light cone originates at the spacetime point (z_a, t_a) and an inverted light zone terminates at (z_b, t_b), see Fig. 2. With $z_b - z_a = M \Delta z$ and $t_b - t_a = N \Delta t$, the edges of the allowable rectangular region are given by $P \equiv \lfloor \frac{N-M}{2} \rfloor$ and $Q \equiv \lceil \frac{N+M}{2} \rceil$, for $N \geq M \geq 0$. Hence,

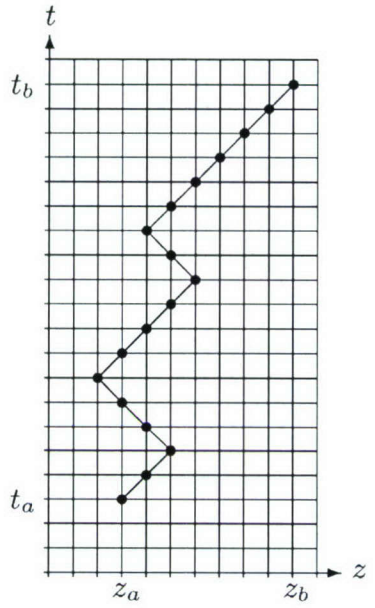


Fig. 1. Example trajectory of a massive relativistic particle starting at location z_a at time t_a and ending at z_b at time t_b. The total number of steps is $N=17$, so the elapsed time is $t=17\Delta t$. The number of steps to the right minus the number to the left is $M=7$, so the net distance traversed is $z=7\Delta z$. The relativistic particle moves at the speed of light $c\equiv\Delta z/\Delta t$. The number of bends is $R=4$.

the paths are the permutations of the set with $N=P+Q$ members ± 1:

$$(\underbrace{1,1,\ldots,1}_{Q\text{ number of }1\text{'s}}\ \underbrace{-1,-1,\ldots,-1}_{P\text{ number of }-1\text{'s}}). \tag{8}$$

The number of permutations is the binomial coefficient:

$$\text{number of paths}=\binom{P+Q}{P}=\binom{P+Q}{Q}. \tag{9}$$

2.3. Spin System Representation

Because the summation (7) occurs on a discrete spacetime lattice, in 1+1 dimensions it is possible to enumerate all the paths originating at point a and ending at point b using $N+1$ spin variables s_i, for $i=0,1,2,\ldots,N$. This is depicted in Fig. 3a for the example relativistic trajectory. Identifying α with s_0 and β with s_N, the summation (7) is equivalent to:

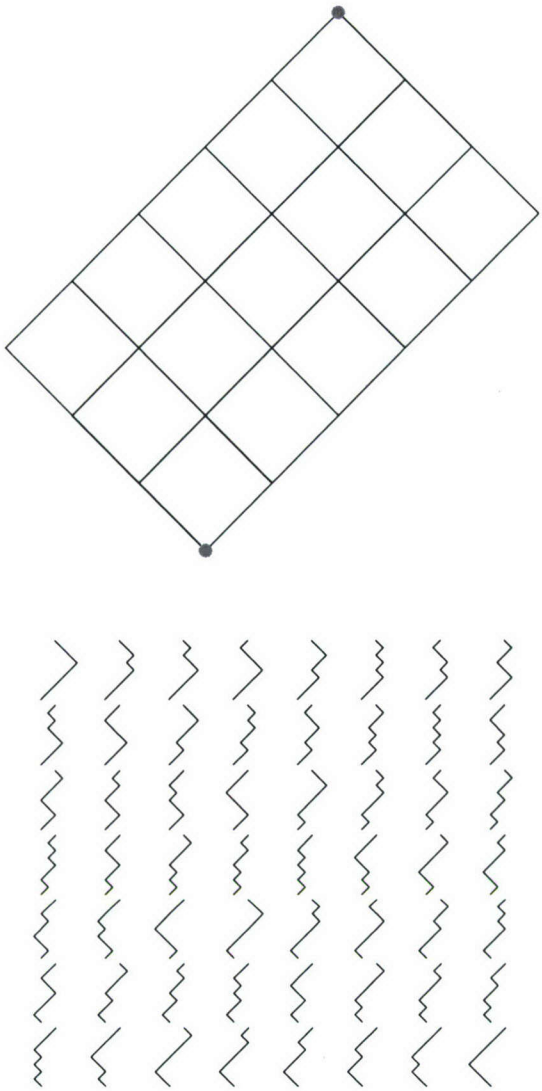


Fig. 2. Example problem in 1+1 dimensions with $N = 8$ and $M = 2$. (Top) Allowable rectangular region of the square lattice within the light cone, with sides of length $Q = \lceil \frac{N+M}{2} \rceil = 5$ and $P = \lfloor \frac{N-M}{2} \rfloor = 3$. (Bottom) Enumeration of all possible paths, the 56 permutations of the set $(1, 1, 1, 1, 1, -1, -1, -1)$.

$$K_{s_0 s_N}(z_a t_a; z_b t_b) = \sum_{\{s_1, \ldots, s_{N-1}\}|\mathrm{M}} \left(i \Delta t \, \frac{mc^2}{\hbar} \right)^R , \tag{10}$$

where the set $\{s_0, \ldots, s_N\}$ specifies a discretized trajectory with a path length constrained by the condition

$$\sum_{i=0}^{N} s_i = \frac{z_b - z_a}{\Delta z} . \tag{11}$$

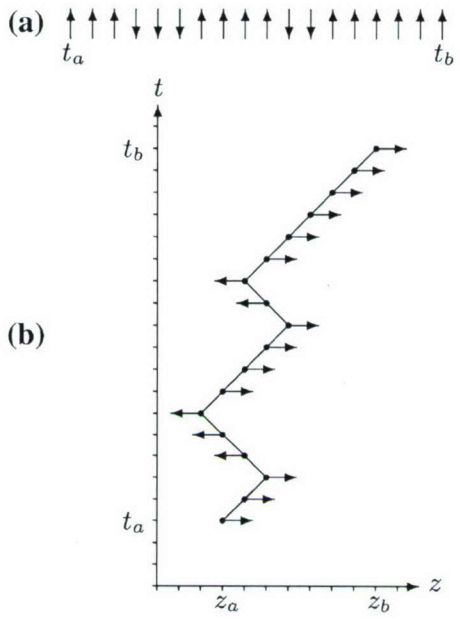


Fig. 3. (a) Spin representation of the trajectory of a massive relativistic particle starting at time t_a and ending at time t_b for $N=7$ and $M=17$. (b) Quantum lattice-gas representation of the same trajectory where the particle is stream-plus (or "spin-up") as it moves to the right and stream-minus (or "spin-down") as it moves to the left. Pre-collision spin orientations are shown.

Since the starting and ending points s_0 ands_N are fixed, they do not appear in the sum on the R.H.S. of (10), and the condition (11) is equivalent to fixing the spin magnetization $M=\sum_i s_i$ of a system of $N+1$ spins.

At the ith step, the particle continues to move straight when $s_i=s_{i+1}$, and it changes direction when $s_i=-s_{i+1}$. As the particle moves (or *streams*) to the right its spin orientation is "spin-up" and as it moves to the left it is "spin-down," as shown in Fig. 3b. Therefore, the following binary value counts the occurrence of a bend at the ith step:

$$\frac{1}{2}(1-s_i s_{i+1})=\begin{cases}0, & \text{no bend}\\ 1, & \text{bend.}\end{cases} \tag{12}$$

Hence, the following sum counts the total number of bends in a path:

$$R=\frac{1}{2}\sum_{i=0}^{N-1}(1-s_i s_{i+1}). \tag{13}$$

With a change of variables

$$\nu \equiv -\frac{1}{2}\log\left(i\Delta t\,\frac{mc^2}{\hbar}\right), \tag{14}$$

(10) can be written as the partition function of an ensemble of spins with nearest neighbor coupling and with fixed total magnetization

$$K_{s_0 s_N} = \sum_{\{s_1,\ldots,s_{N-1}\}} \delta\left(M, \sum_i s_i\right) e^{-\nu\sum_{i=0}^{N-1}(1-s_i s_{i+1})}, \tag{15}$$

where the Kronecker delta $\delta(a,b)=1$ for $a=b$ and $\delta(a,b)=0$ for $a\neq b$. We may write the Kronecker delta as follows:

$$\delta\left(M, \sum_{i=0}^{N} s_i\right) = \frac{1}{2N}\sum_{n=-N}^{N-1} e^{i\frac{2\pi n}{N}(M-\sum_i s_i)}, \tag{16}$$

since M and $\sum_{i=0}^{N} s_i$ are integers. Then inserting (16) into (15) gives

$$K_{s_0 s_N} = \frac{1}{2N}\sum_{n=-N}^{N-1} e^{i\left(\frac{2\pi n}{N}\right)M} \sum_{\{s_1,\ldots,s_{N-1}\}} e^{-i\left(\frac{2\pi n}{N}\right)\sum_{i=0}^{N} s_i - \nu\sum_{i=0}^{N-1}(1-s_i s_{i+1})}. \tag{17}$$

Now since $\sum_{i=0}^{N} s_i = \frac{1}{2}(s_0+s_N)+\frac{1}{2}\sum_{i=0}^{N-1}(s_i+s_{i+1})$, we pull down the summation in the argument of the exponential to form the following product:

$$\begin{aligned} K_{s_0 s_N} = {} & \frac{1}{2N}\sum_{n=-N}^{N-1} e^{i\left(\frac{2\pi n}{N}\right)M} e^{-i\pi\left(\frac{n}{N}\right)(s_0+s_N)} \\ & \times \sum_{\{s_1,\ldots,s_{N-1}\}} \prod_{i=0}^{N-1} e^{-i\pi\left(\frac{n}{N}\right)(s_i+s_{i+1})-\nu(1-s_i s_{i+1})}. \end{aligned} \tag{18}$$

The components of a unitary transfer matrix $\mathscr{U}$ are defined as

$$\mathscr{U}_{s_i,s_{i+1}} \equiv e^{-\nu(1-s_i s_{i+1})-i\pi\left(\frac{n}{N}\right)(s_i+s_{i+1})}, \tag{19}$$

so that (18) becomes

$$K_{s_0 s_N} = \frac{1}{2N}\sum_{n=-N}^{N-1} e^{i\left(\frac{2\pi n}{N}\right)M} e^{-i\pi\left(\frac{n}{N}\right)(s_0+s_N)} \mathscr{L}_{s_0 s_N}, \tag{20}$$

where we have defined

$$\mathscr{Z}_{s_0 s_N} \equiv \sum_{s_1=\pm 1} \cdots \sum_{s_{N-1}=\pm 1} \mathscr{U}_{s_0,s_1} \mathscr{U}_{s_1,s_2} \cdots \mathscr{U}_{s_{N-1},s_N}. \tag{21}$$

The matrix form of (19) is

$$\mathscr{U} = \begin{pmatrix} \mathscr{U}_{1,1} & \mathscr{U}_{-1,1} \\ \mathscr{U}_{1,-1} & \mathscr{U}_{-1,-1} \end{pmatrix} = \begin{pmatrix} e^{-i2\pi n/N} & e^{-2\nu} \\ e^{-2\nu} & e^{i2\pi n/N} \end{pmatrix} \tag{22}$$

and so (21) becomes simply an $N-1$ fold matrix multiplication of $\mathscr{U}$:

$$\mathscr{Z} = \begin{pmatrix} \mathscr{Z}_{1,1} & \mathscr{Z}_{-1,1} \\ \mathscr{Z}_{1,-1} & \mathscr{Z}_{-1,-1} \end{pmatrix} = \begin{pmatrix} e^{-i2\pi n/N} & e^{-2\nu} \\ e^{-2\nu} & e^{i2\pi n/N} \end{pmatrix}^N. \tag{23}$$

2.4. Quantum Lattice Gas Representation

To simplify (20), we acknowledge the discretized nature of the space-time lattice by using the following variable transformations:

$$\frac{2\pi n}{N} = \frac{p_n \Delta z}{\hbar} \qquad z = z_b - z_a = \Delta z M \qquad t = t_b - t_a = \Delta t N, \tag{24}$$

where p_n is the discretized momentum. From (14), the off-diagonal components of (22) are $e^{-2\nu} = i\Delta t \frac{mc^2}{\hbar}$, and the transfer matrix (22) becomes

$$\mathscr{U} = \begin{pmatrix} 1 - i\frac{p_n \Delta z}{\hbar} + \mathscr{O}(\Delta z^2) & i\Delta t \frac{mc^2}{\hbar} \\ i\Delta t \frac{mc^2}{\hbar} & 1 + i\frac{p_n \Delta z}{\hbar} + \mathscr{O}(\Delta z^2) \end{pmatrix}. \tag{25}$$

Our basic approach to formulate the space-time translation degrees of freedom in terms of the spin variables is to partition (25) into a product of a temporal matrix and a spatial matrix

$$\mathscr{U} = \mathscr{C} S, \tag{26}$$

where the temporal matrix is called the *collision matrix*

$$\mathscr{C} = e^{i\sigma_x \frac{mc^2 \Delta t}{\hbar}} \tag{27a}$$

$$= \begin{pmatrix} \cos\left(\frac{mc^2 \Delta t}{\hbar}\right) & i \sin\left(\frac{mc^2 \Delta t}{\hbar}\right) \\ i \sin\left(\frac{mc^2 \Delta t}{\hbar}\right) & \cos\left(\frac{mc^2 \Delta t}{\hbar}\right) \end{pmatrix} \tag{27b}$$

$$= \begin{pmatrix} 1 & i\Delta t \frac{mc^2}{\hbar} \\ i\Delta t \frac{mc^2}{\hbar} & 1 \end{pmatrix} + \mathscr{O}(\Delta t^2), \tag{27c}$$

and the spatial matrix is called the *stream matrix*

$$\mathcal{S} = e^{-i\sigma_z \frac{p_n \Delta z}{\hbar}} \tag{28a}$$

$$= \begin{pmatrix} e^{-i\frac{p_n \Delta z}{\hbar}} & 0 \\ 0 & e^{i\frac{p_n \Delta z}{\hbar}} \end{pmatrix} \tag{28b}$$

$$= \begin{pmatrix} 1 - i\frac{p_n \Delta z}{\hbar} & 0 \\ 0 & 1 + i\frac{p_n \Delta z}{\hbar} \end{pmatrix} + \mathcal{O}(\Delta z^2). \tag{28c}$$

The decomposition of $\mathcal{U}$ given in (26) is a *quantum lattice gas representation* of the transfer matrix. In this construction with $\Delta z = c\Delta t$ where the particles stream through space at a fixed unit velocity equal to the speed of light, (23) becomes

$$\mathcal{L} = (\mathcal{C}\mathcal{S})^N \tag{29a}$$

$$= e^{i(\sigma_x mc^2 - \sigma_z p_n c)N\Delta t/\hbar} + \mathcal{O}(\Delta t^2). \tag{29b}$$

Inserting (29b) into (20), the quantum lattice-gas kernel becomes

$$K_{s_0 s_N} = \frac{1}{2N} \sum_{n=-N}^{N-1} e^{i\frac{p_n z}{\hbar}} e^{-\frac{ip_n \Delta z}{2\hbar}(s_0 + s_N)} \left[e^{i(\sigma_x mc^2 - \sigma_z p_n c)N\Delta t/\hbar} \right]_{s_0 s_N}. \tag{30}$$

As N approaches infinity in the continuum limit, we can neglect $\Delta z = c\Delta t$ compared with $t = N\Delta t$ (that is, $N \gg c$), so we have

$$K_{s_0 s_N} = \frac{1}{2N} \sum_{n=-N}^{N-1} e^{i\frac{p_n z}{\hbar}} \left[e^{i(\sigma_x mc^2 - \sigma_z p_n c)N\Delta t/\hbar} \right]_{s_0 s_N}. \tag{31}$$

In the continuum limit, the summation goes over to an integral ($\frac{1}{N}\sum_n \to \frac{dz}{h}\int dp$) and so we have

$$K_{\alpha\beta}(z,t) \equiv \lim_{N\to\infty} K_{s_0 s_N} \tag{32a}$$

$$= \frac{dz}{2h} \int_{-\infty}^{\infty} dp\, e^{i\frac{pz}{\hbar}} \left[e^{i(\sigma_x mc^2 - \sigma_z pc)t/\hbar} \right]_{\alpha\beta}. \tag{32b}$$

Our result (32) is equivalent to the result found by Jacobson and Schulman [3]. However our derivation, based on the quantum lattice gas representation (26), is simpler. Furthermore, the quantum lattice gas representation allows us to generalize the derivation to obtain the Dirac equation in 3+1 dimensions.

3. 3+1 DIMENSIONS

3.1. Feynman Path Integral Representation of Non-relativistic Quantum Mechanics

The probability amplitude that a quantum particle at position $\vec{r}_a = (x_a, y_a, z_a)$ at time t_a will transfer to a new position $\vec{r}_b = (x_b, y_b, z_b)$ at time t_b is given by the following path integral:

$$K(\vec{r}_a t_a; \vec{r}_b t_b) = \int \mathcal{D}[\vec{r}(t)] e^{i\frac{S[\vec{r}(t)]}{\hbar}}, \tag{33}$$

where $\int \mathcal{D}[\vec{r}(t)]$ denotes integration over all trajectories $\vec{r}(t)$ for which $\vec{r}(t_a) = \vec{r}_a$ and $\vec{r}(t_b) = \vec{r}_b$, and where the increase of the action

$$S = \int_{t_a}^{t_b} dt\, L[\dot{\vec{r}}(t), \vec{r}(t)], \tag{34}$$

along a trajectory $\vec{r}(t)$ is determined using the classical Lagrangian L. Equation (33) is a solution of the quantum mechanical wave equation.[6]

3.2. Path summation Representation of Relativistic Quantum Mechanics

We introduce a discrete representation of the Feynman path integral principle on a body-centered cubic spatial lattice with discrete time steps to compute (33) for a relativistic quantum particle:

$$K_{\vec{\alpha}\vec{\beta}}(\vec{r}_a t_a; \vec{r}_b t_b) = \lim_{N\to\infty} \prod_{w=1}^{D} \sum_{R_w \geq 0} \Phi_{\vec{\alpha}\vec{\beta}}(R_w) \left(i \frac{mc^2 \Delta t}{D\hbar} \right)^{R_w}, \tag{35}$$

where $\Delta t \equiv \frac{t_b - t_a}{N}$, where $\vec{\alpha}$ and $\vec{\beta}$ represented by initial and final spin vectors of the form $(s_x, s_y, s_z) = (\pm 1, \pm 1, \pm 1)$, where $K_{\vec{\alpha}\vec{\beta}}$ is the kernel of a 4-spinor amplitude field, where the length of each step is $\Delta r \equiv c\, \Delta t$, where c is the speed of light, where m is the mass of the quantum particle, where R_w is the number of bends in a path counted by projecting the path along the $\hat{w}$ axis, and where $\Phi_{\vec{\alpha}\vec{\beta}}(R_w)$ is the number of allowable paths with R_w bends and N steps along the $\hat{w}$ axis, and where D is the number of spatial dimensions. The separation distance between the starting and ending points is:

$$\vec{r}_b - \vec{r}_a = \hat{r}\sqrt{(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2}. \tag{36}$$

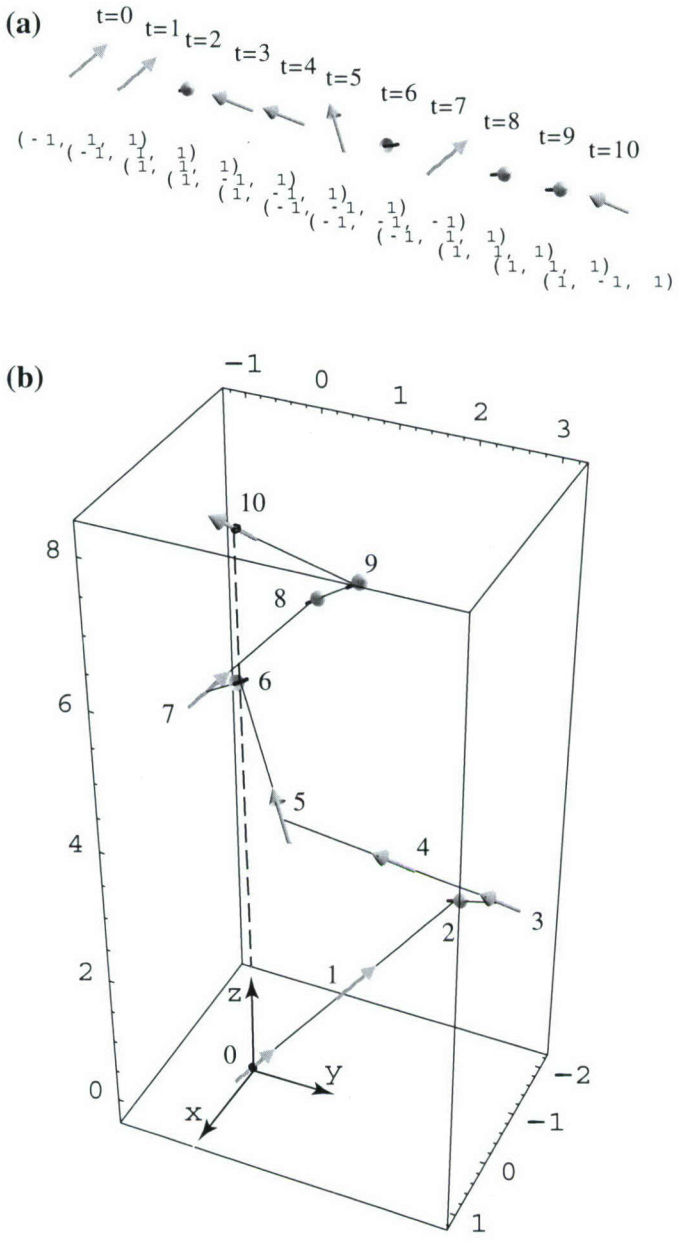


Fig. 4. (a) Spin representation of the trajectory of a massive relativistic particle. The components of each spin vector are shown. (b) Quantum lattice-gas representation of an example trajectory of a relativistic particle (solid lines). The coordinate system is chosen so that the final displacement (dashed line) is along the $\hat{z}$-axis direction: the particle starts at location $\vec{r}_a = \vec{0}$ at time $t_a = 0$ (labeled "0") and ends at $\vec{r}_b = (8\Delta z)\hat{z}$ at time $t_b = 10\Delta t$ (labeled "10"). The total number of steps is $N = 10$. There are 3 bends along $\hat{x}$ at $t = 2, t = 5$, and $t = 8$; 3 bends along $\hat{y}$ at $t = 3, t = 7$, and $t = 10$; and 2 bends along $\hat{z}$ at $t = 6\Delta t$ and $t = 7\Delta t$. The particle moves at the speed of light $c = \Delta r/\Delta t$ along three cubic lattice directions simultaneously. The spin at each step specifies the direction of motion for that step. Post-collision spin orientations are shown.

An example three dimensional relativistic trajectory with 3 bends along $\hat{x}$ and $\hat{y}$, and 2 bends along $\hat{z}$ is depicted in Fig. 4 where a bend occurs when the particle changes its direction of motion along any axis. Equation (35) reduces to the Feynman chessboard formula (7) when $D=1$.

3.3. Spin System Representation

With three spatial dimensions, it is possible to enumerate all the paths originating at point $\vec{a}$ and ending at point $\vec{b}$ using $N+1$ spin variables $\vec{s}_n=(s_{xn}, s_{yn}, s_{zn})$, for $n=0, 1, 2, \ldots, N$, and where $s_{wn}=\pm 1$ for $w=x, y, z$. There are eight distinct spin vectors, one for each direction along a diagonal of the body centered cubic (bcc) lattice $\vec{s}_n=(\pm 1, \pm 1, \pm 1)$. This is depicted in Fig. 5. Note that $\vec{s}_n$ is a discrete 3-vector whereas the spin vector $\vec{\sigma}=(\sigma_x, \sigma_y, \sigma_z)$ is a 3-vector of 2×2 Pauli matrices. We shall see in the following exactly how and why $\vec{s}$ and $\vec{\sigma}$ are related. The result is that a Dirac particle quantum mechanically moves along a diagonal of the bcc lattice with its spin vector $\vec{\sigma}$ "parallel" to its direction of motion $c\vec{s}$.

Identifying $\vec{\alpha}$ with $\vec{s}_0$ and $\vec{\beta}$ with $\vec{s}_N$, the summation (35) is equivalent to:

$$K_{\vec{s}_0\vec{s}_N}(\vec{r}_a t_a; \vec{r}_b t_b)=\sum_{\{\vec{s}_1,\ldots,\vec{s}_{N-1}\}|\vec{M}}\left(i\frac{mc^2\Delta t}{D\hbar}\right)^R, \tag{37}$$

where $R=\sum_{w=1}^{D} R_w$, where R_w is a bend along the $\hat{w}$ direction, and where the set $\{\vec{s}_0, \ldots, \vec{s}_N\}$ specifies a discretized trajectory with a path length constrained by the condition

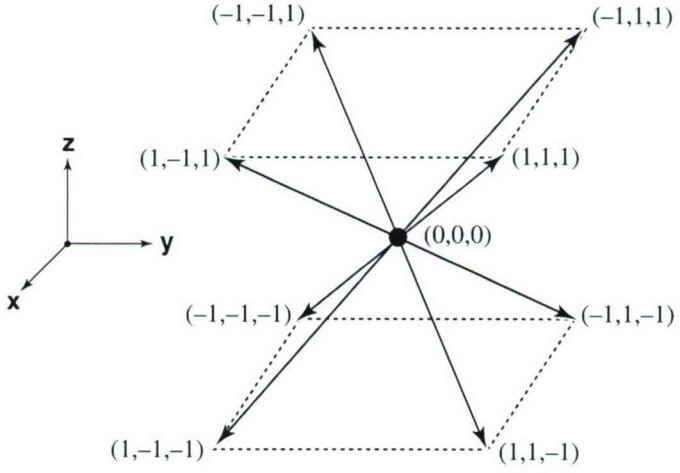


Fig. 5. Lattice diagonal directions for a body-centered cubic spatial lattice for a quantum particle starting at the origin $\vec{r}_a=(0, 0, 0)$.

$$\sum_{n=0}^{N} \vec{s}_n = \left(\frac{x_b - x_a}{\Delta x}, \frac{y_b - y_a}{\Delta y}, \frac{z_b - z_a}{\Delta z} \right). \tag{38}$$

Since the starting and ending points $\vec{s}_0$ and $\vec{s}_N$ are fixed, they do not appear in the sum on the R.H.S. of (37), and the condition (38) is equivalent to fixing the spin magnetization $\vec{M} = \sum_n \vec{s}_n$ of a system of $N+1$ spins.

At the nth step, the particle continues to move straight when $\vec{s}_n \cdot \vec{s}_{n+1} = D$, reverses direction along on orthogonal axis when $\vec{s}_n \cdot \vec{s}_{n+1} = 1$, reverses directions along two orthogonal axes when $\vec{s}_n \cdot \vec{s}_{n+1} = -1$, and completely reverses (flips) direction when $\vec{s}_n \cdot \vec{s}_{n+1} = -D$. The following sum counts the occurrence of a bend along the $\hat{w}$ axis at the nth step:

$$\frac{1}{2}[1 - s_{wn} s_{w(n+1)}] = \begin{cases} 0, & \text{no bend along } \hat{w} \\ 1, & \text{bend along } \hat{w}. \end{cases} \tag{39}$$

In turn, the following sum counts the total number of bends in a path between $\vec{r}_a$ and $\vec{r}_b$:

$$R = \frac{1}{2} \sum_{w=1}^{D} \sum_{n=0}^{N-1} [1 - s_{wn} s_{w(n+1)}]. \tag{40}$$

With a simple generalizational of the change of variables (14) used in the one-dimensional case

$$\nu \equiv -\frac{1}{2} \log \left(i \frac{mc^2 \Delta t}{D\hbar} \right), \tag{41}$$

(37) can be written as the partition function of an ensemble of spins with nearest neighbor coupling and with fixed total magnetization $\vec{M} = (M_x, M_y, M_z)$:

$$K_{\vec{s}_0 \vec{s}_N} = \sum_{\{\vec{s}_1, \ldots, \vec{s}_{N-1}\}} \delta \left(\vec{M}, \sum_n \vec{s}_n \right) e^{-\nu \sum_{n=0}^{N-1} [D - \vec{s}_n \cdot \vec{s}_{n+1}]}, \tag{42}$$

where the Kronecker delta is:

$$\begin{aligned} \delta \left(\vec{M}, \sum_{n=0}^{N} \vec{s}_n \right) &= \prod_{w=1}^{D} \delta \left(M_w, \sum_n s_{wn} \right) \\ &= \delta \left(M_x, \sum_n s_{xn} \right) \delta \left(M_y, \sum_n s_{yn} \right) \delta \left(M_z, \sum_n s_{zn} \right), \end{aligned} \tag{43}$$

where

$$\delta\left(M_x, \sum_n s_{xn}\right) = \frac{1}{2N} \sum_{n_x=-N}^{N-1} e^{i\frac{2\pi n_x}{N}(M_x - \sum_n s_{xn})} \tag{44a}$$

$$\delta\left(M_y, \sum_i s_{yn}\right) = \frac{1}{2N} \sum_{n_y=-N}^{N-1} e^{i\frac{2\pi n_y}{N}(M_y - \sum_n s_{yn})} \tag{44b}$$

$$\delta\left(M_z, \sum_i s_{zn}\right) = \frac{1}{2N} \sum_{n_z=-N}^{N-1} e^{i\frac{2\pi n_z}{N}(M_z - \sum_n s_{zn})}. \tag{44c}$$

The total magnetization is related to the total displacement vector as follows:

$$\vec{x} \equiv (M_x \Delta x, M_y \Delta y, M_z \Delta z). \tag{45}$$

Furthermore, the wave number is defined as follows:

$$\vec{k}_{\vec{n}} \equiv \frac{2\pi}{N}\left(\frac{n_x}{\Delta x}, \frac{n_y}{\Delta y}, \frac{n_z}{\Delta z}\right). \tag{46}$$

Using (45) and (46), the Kronecker delta (43) becomes

$$\delta\left(\vec{M}, \sum_{n=0}^{N} \vec{s}_n\right) = \frac{1}{(2N)^3} \sum_{\vec{n}=(n_x,n_y,n_z)} e^{i\vec{k}_{\vec{n}}\cdot\vec{x} - i\Delta r \sum_n \vec{k}_{\vec{n}}\cdot\vec{s}_n}. \tag{47}$$

Then inserting (47) into (42) gives

$$K_{\vec{s}_0\vec{s}_N} = \frac{1}{(2N)^3} \sum_{\vec{n}} e^{i\vec{k}_{\vec{n}}\cdot\vec{x}} \sum_{\{\vec{s}_1,\ldots,\vec{s}_{N-1}\}} e^{-i\Delta r \sum_{n=0}^{N} \vec{k}_{\vec{n}}\cdot\vec{s}_n - \nu \sum_{n=0}^{N-1}[D - \vec{s}_n\cdot\vec{s}_{n+1}]}. \tag{48}$$

We can proceed from (48) using the identity $\sum_{n=0}^{N} \vec{s}_n = \frac{1}{2}(\vec{s}_0 + \vec{s}_N) + \frac{1}{2}\sum_{n=0}^{N-1}(\vec{s}_n + \vec{s}_{n+1})$, which allows us to pull down the summation in the argument of the exponential to form the following product:

$$K_{\vec{s}_0\vec{s}_N} = \frac{1}{(2N)^3}\sum_{\vec{n}} e^{i\vec{k}_{\vec{n}}\cdot\vec{x}} e^{-\frac{i}{2}\Delta r\vec{k}_{\vec{n}}\cdot(\vec{s}_0+\vec{s}_N)}$$

$$\times \sum_{\{\vec{s}_1,\ldots,\vec{s}_{N-1}\}} \prod_{n=0}^{N-1} e^{-\frac{i}{2}\Delta r\vec{k}_{\vec{n}}\cdot(\vec{s}_n+\vec{s}_{n+1})-\nu[D-\vec{s}_n\cdot\vec{s}_{n+1}]} \tag{49a}$$

$$= \frac{1}{(2N)^3}\sum_{\vec{n}} e^{i\vec{k}_{\vec{n}}\cdot\vec{x}} e^{-\frac{i}{2}\Delta r\vec{k}_{\vec{n}}\cdot(\vec{s}_0+\vec{s}_N)} \prod_{w=1}^{D}$$

$$\times \sum_{\{s_{w1},\ldots,s_{w(N-1)}\}} \prod_{n=0}^{N-1} e^{-\frac{i}{2}\Delta r k_{nw}(s_{wn}+s_{w(n+1)})-\nu(1-s_{wn}s_{w(n+1)})}. \tag{49b}$$

As in the one-dimensional case, here again we use a unitary transfer matrix. The components of the transfer matrix $\mathscr{U}$ for the cartesian axis $\hat{w}$ are defined as

$$\mathscr{U}_{s_{wn},s_{w(n+1)}} \equiv e^{-\frac{i}{2}\Delta r k_{nw}(s_{wn}+s_{w(n+1)})-\nu(1-s_{wn}s_{w(n+1)})}, \tag{50}$$

then (49b) becomes

$$K_{\vec{s}_0\vec{s}_N} = \frac{1}{(2N)^3}\sum_{\vec{n}} e^{i\vec{k}_{\vec{n}}\cdot\vec{x}} e^{-\frac{i}{2}\Delta r\vec{k}_{\vec{n}}\cdot(\vec{s}_0+\vec{s}_N)} \mathscr{L}_{\vec{s}_0\vec{s}_N}, \tag{51}$$

where we have defined

$$\mathscr{L}_{\vec{s}_0\vec{s}_N} \equiv \prod_{w=1}^{D} \sum_{\{s_{w1},\ldots,s_{w(N-1)}\}} \mathscr{U}_{s_{w0},s_{w1}}\mathscr{U}_{s_{w1},s_{w2}}\cdots\mathscr{U}_{s_{w(N-1)},s_{wN}}. \tag{52}$$

The generic sum over all possible spin components in (52) can be rewritten explicitly as $N-1$ binary sums:

$$\mathscr{L}_{\vec{s}_0\vec{s}_N} = \prod_{w=1}^{D} \sum_{s_{w1}=\pm1} \cdots \sum_{s_{w(N-1)}=\pm1} \mathscr{U}_{s_{w0},s_{w1}}\mathscr{U}_{s_{w1},s_{w2}}\cdots\mathscr{U}_{s_{w(N-1)},s_{wN}}. \tag{53}$$

The form of (53) suggests that we can write the transfer matrix (50) for the $\hat{w}$ direction as a 2×2 matrix where the binary sums represent matrix multiplication, just as in the 1+1 dimensional case. However, in accordance with the product $\prod_w$ in (53), the sums over the orthogonal cartesian axes $\hat{w}$ directions must be kept independent. The only alternative when enumerating the four spin components, $s_{wn}=\pm1$ and $s_{w(n+1)}=\pm1$, is to introduce D number of new token variables, say μ_w, to keep the directional sums independent of each other. The kinetic term

$e^{\frac{i}{2}\Delta r k_{nw}(s_{wn}+s_{w(n+1)})}$ in (50) contributes to only the diagonal (non-reversal) components in (55) when $s_{wn} = s_{w(n+1)}$, and contributes 0 to the non-diagonal (reversal) components when $s_{wn} = -s_{w(n+1)}$. For its diagonal contribution, we take the local translation operator

$$e^{-\frac{i}{2}\Delta r k_{wn}(s_{wn}+s_{w(n+1)})} \to \begin{cases} e^{-i\Delta r k_{wn}\mu_w} & s_{wn}=1 \\ & s_{w(n+1)}=1, \\ e^{i\Delta r k_{wn}\mu_w} & s_{wn}=-1 \\ & s_{w(n+1)}=-1. \end{cases} \tag{54}$$

The mass term $e^{-\nu(1-s_{wn}s_{w(n+1)})}$ in (50) contributes 0 to the diagonal (non-reversal) components in (55) when $s_{wn} = s_{w(n+1)}$ and contributes $e^{-2\nu}$ to the non-diagonal (reversal) components when $s_{wn} = -s_{w(n+1)}$. Hence, (50) is written in 2×2 matrix form as follows:

$$\mathscr{U}_w = \begin{pmatrix} \mathscr{U}_{1,1} & \mathscr{U}_{-1,1} \\ \mathscr{U}_{1,-1} & \mathscr{U}_{-1,-1} \end{pmatrix} = \begin{pmatrix} e^{-i\Delta r k_{wn}\mu_w} & e^{-2\nu} \\ e^{-2\nu} & e^{i\Delta r k_{wn}\mu_w} \end{pmatrix}. \tag{55}$$

Notice we have no need to employ a token variable that depends on w for the off-diagonal components of (55) since the term $e^{-2\nu}$ for each $\hat{w}$ direction contributes equally to the mass of the Dirac particle. That is, since bends in the $\hat{x}$, $\hat{y}$, or $\hat{z}$ directions equally contribute to the particle's mass, all the $e^{-2\nu}$ components in (51) may add together. So we can omit using another token variable for the off-diagonal terms.

With (55), we see that (53) is simply an $N-1$ fold matrix multiplication of $\mathscr{U}$:

$$\mathscr{L} = \begin{pmatrix} \mathscr{L}_{1,1} & \mathscr{L}_{-1,1} \\ \mathscr{L}_{1,-1} & \mathscr{L}_{-1,-1} \end{pmatrix} = \prod_{w=1}^{D} \begin{pmatrix} e^{-i\Delta r k_{wn}\mu_w} & e^{-2\nu} \\ e^{-2\nu} & e^{i\Delta r k_{wn}\mu_w} \end{pmatrix}^N. \tag{56}$$

At this point, we have not yet determined a specific form of the object μ_w. This will be determined in the next section.

3.4. Fourier Transform of the Dirac Propagator

Just as we had previously in the one-dimensional case, we can rewrite (51) using momentum variables. Along with (45) and (46), we have the following variable transformations:

$$p_{wn} \equiv \hbar k_{wn} \qquad t = t_b - t_a = \Delta t N. \tag{57}$$

From our change of variables (14), the off-diagonal components of (55) are $e^{-2\nu} = i\Delta t \frac{mc^2}{D\hbar}$, and the transfer matrix (55) becomes

$$\mathscr{U}_w = \begin{pmatrix} e^{-i\Delta r \frac{p_{wn}\mu_w}{\hbar}} & i\Delta t \frac{mc^2}{D\hbar} \\ i\Delta t \frac{mc^2}{D\hbar} & e^{i\Delta r \frac{p_{wn}\mu_w}{\hbar}} \end{pmatrix} \tag{58a}$$

$$= e^{-i\sigma_z \mu_w p_{wn} \Delta r/\hbar} + i\sigma_x \frac{mc^2}{D\hbar}\Delta t, \tag{58b}$$

where σ_x and σ_z are Pauli matrices.

As in the 1+1 dimensional case, our basic approach here is to represent the spacetime translation degrees of freedom in terms of the spin variables by writing (58b) as a product of a collision matrix and a stream matrix

$$\mathscr{U}_w = e^{-i\sigma_z \mu_w p_{wn} \Delta r/\hbar} e^{i\sigma_x \frac{mc^2}{D\hbar}\Delta t} \tag{59a}$$

$$= \mathscr{C}\mathscr{S}_w, \tag{59b}$$

where in the 3+1 dimensional case we now have

$$\mathscr{C} \equiv e^{i\sigma_x \frac{mc^2}{D\hbar}\Delta t}, \tag{60}$$

and

$$\mathscr{S}_w \equiv e^{-i\sigma_z \mu_w p_{wn} \Delta r/\hbar}. \tag{61}$$

Inserting (59) into (56), we have

$$\mathscr{Z} = \prod_{w=1}^{D} (\mathscr{C}\mathscr{S}_w)^N \tag{62a}$$

$$= \prod_{w=1}^{D} e^{-\frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \sigma_x \frac{mc^2}{D}\right) N\Delta t}. \tag{62b}$$

Now with (57), we have

$$\mathscr{Z}(t) = \prod_{w=1}^{D} e^{-\frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \sigma_x \frac{mc^2}{D}\right) t}. \tag{63}$$

We impose a physical constraint on $\mathscr{Z}(t)$ in (63) that the quantum mechanical propagation of a Dirac particle be factorable in time as follows:

$$\mathscr{Z}(t_1 + t_2) = \mathscr{Z}(t_2)\mathscr{Z}(t_1). \tag{64}$$

Let us square the propagator by choosing $t_1 = t$ and $t_2 = t$ so by inserting (63) into (64) we have

$$\prod_{w=1}^{D} e^{-\frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \sigma_x \frac{mc^2}{D}\right)2t} = \left(\prod_{w=1}^{D} e^{-\frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \sigma_x \frac{mc^2}{D}\right)t}\right) \times \left(\prod_{w=1}^{D} e^{-\frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \sigma_x \frac{mc^2}{D}\right)t}\right)$$
$$= \left(1 - \sum_{w=1}^{D} \frac{i}{\hbar}\left(\sigma_z \mu_w p_{wn} c - \frac{\sigma_x mc^2}{D}\right)t + \cdots\right) \times \left(1 - \sum_{w'=1}^{D} \frac{i}{\hbar}\left(\sigma_z \mu_{w'} p_{w'n} c - \frac{\sigma_x mc^2}{D}\right)t + \cdots\right), \tag{65}$$

where in the Taylor expansions on the R.H.S we must keep all high order terms since, in the continuum limit, $t = \lim_{N\to\infty} N\Delta t$ is not a small parameter. The only way for the L.H.S. to equal the R.H.S. is for the all the cross-terms to vanish. Now, all the cross-terms between the momentum and time components automatically vanish because of the anti-commutation of the Pauli matrices σ_z and σ_x. However, for all the cross-terms between the momentum components to vanish requires our token variables μ_w be anti-commuting Grassman variables:

$$\mu_i \mu_j + \mu_j \mu_i = 0, \tag{66}$$

for $1 \le i \le D$ and $1 \le j \le D$. We are free to represent the μ_w variables in terms of Pauli matrices by choosing

$$\mu_w \to \sigma_w. \tag{67}$$

In this representation, $\mathscr{C}$ and $\mathscr{S}$ become 4×4 matrices. The collision operator is

$$\mathscr{C} = \prod_{w=1}^{D} e^{i\sigma_x \otimes \mathbf{1} \frac{mc^2 \Delta t}{D\hbar}} = e^{i\sigma_x \otimes \mathbf{1} \frac{mc^2 \Delta t}{\hbar}} \tag{68a}$$

$$= \begin{pmatrix} \cos\left(\Delta t \frac{mc^2}{\hbar}\right) & i \sin\left(\Delta t \frac{mc^2}{\hbar}\right) \\ i \sin\left(\Delta t \frac{mc^2}{\hbar}\right) & \cos\left(\Delta t \frac{mc^2}{\hbar}\right) \end{pmatrix} \otimes \mathbf{1} \tag{68b}$$

$$= \begin{pmatrix} \mathbf{1} & i\Delta t \frac{mc^2}{\hbar}\mathbf{1} \\ i\Delta t \frac{mc^2}{\hbar}\mathbf{1} & \mathbf{1} \end{pmatrix} + \mathcal{O}(\Delta t^2) \tag{68c}$$

$$= \begin{pmatrix} 1 & 0 & i\Delta t\,\frac{mc^2}{\hbar} & 0 \\ 0 & 1 & 0 & i\Delta t\,\frac{mc^2}{\hbar} \\ i\Delta t\,\frac{mc^2}{\hbar} & 0 & 1 & 0 \\ 0 & i\Delta t\,\frac{mc^2}{\hbar} & 0 & 1 \end{pmatrix}, \tag{68d}$$

and the stream operator is

$$\mathcal{S} = \prod_{w=1}^{D} e^{-\frac{i}{\hbar}\sigma_z \otimes \sigma_w p_{wn} c\Delta t} = e^{-\frac{i}{\hbar}\sigma_z \otimes \vec{\sigma}\cdot\vec{p}_{\vec{n}} c\Delta t} = \begin{pmatrix} e^{-\frac{i}{\hbar}\sigma\cdot\vec{p}_{\vec{n}} c\Delta t} & 0 \\ 0 & e^{\frac{i}{\hbar}\vec{\sigma}\cdot\vec{p}_{\vec{n}} c\Delta t} \end{pmatrix} \tag{69a}$$

$$= \begin{pmatrix} \mathbf{1} - \frac{i\Delta r}{\hbar}\vec{\sigma}\cdot\vec{p}_{\vec{n}} & 0 \\ 0 & \mathbf{1} + \frac{i\Delta r}{\hbar}\vec{\sigma}\cdot\vec{p}_{\vec{n}} \end{pmatrix} + \mathcal{O}(\Delta r^2) \tag{69b}$$

$$= \begin{pmatrix} 1 - \frac{i\Delta r p_{zn}}{\hbar} & -\frac{i\Delta r(p_{xn} - ip_{yn})}{\hbar} & 0 & 0 \\ -\frac{i\Delta r(p_{xn} + ip_{yn})}{\hbar} & 1 + \frac{i\Delta r p_{zn}}{\hbar} & 0 & 0 \\ 0 & 0 & 1 + \frac{i\Delta r p_{zn}}{\hbar} & \frac{i\Delta r(p_{xn} - ip_{yn})}{\hbar} \\ 0 & 0 & \frac{i\Delta r(p_{xn} + ip_{yn})}{\hbar} & 1 - \frac{i\Delta r p_{zn}}{\hbar} \end{pmatrix}. \tag{69c}$$

Equation (69) is not diagonal and therefore it is a quantum mechanical stream operator; a single particle moves from one place to multiple places in one time Δt. Even a wave function of a massless particle will undergo dispersion. So a quantum particle does not have a single trajectory since its position continually spreads out as it moves. This is compatible with the Heisenberg uncertainty principle.[4]

Finally, using (68a) and (69a), and neglecting terms proportional to $\frac{\Delta r}{N} \ll \Delta t$, the kernal (62b) becomes the following three dimensional

[4] Niels Bohr objected to Feynman's talk at the 1948 Pocono conference for the reason that the classical paths Feynman was enumerating were not consistent with the Heisenberg uncertainty principle.[7] With our 3+1 dimensional solution, we see from (69) that Bohr's objective could have been resolved by answering that the path integral formalism *does* accommodate the notion of path spreading, consistent with the Heisenberg uncertainty principle. However, Feynman used the viewpoint referred to as "calculate and shut-up" in response to Bohr (the quote now attributed to David Mermin). By that time although Feynman had worked out relativistic QED in 3+1 dimensions, he had only worked out discrete spacetime pictures of the particle motion for the 1+1 dimensional square lattice, where the translation is classical.

discrete Fourier transform:

$$K_{\vec{\sigma}_0\vec{\sigma}_N} = \frac{1}{(2N)^3} \sum_{\vec{n}} e^{i\frac{\vec{p}_{\vec{n}}\cdot\vec{x}}{\hbar}} \left[e^{-\frac{i}{\hbar}\left(\sigma_z \otimes \sigma \cdot \vec{p}_{\vec{n}} c - \sigma_x \otimes \mathbf{1} mc^2\right)t} \right]_{\vec{\sigma}_0\vec{\sigma}_N}. \tag{70}$$

In the continuum limit, we recover the kernel

$$K_{\vec{\alpha}\vec{\beta}} = \frac{dx^3}{(2h)^3} \int dp^3 e^{i\frac{\vec{p}\cdot\vec{x}}{\hbar}} \left[e^{-\frac{i}{\hbar}H_{\text{Dirac}}t} \right]_{\vec{\alpha}\vec{\beta}}, \tag{71}$$

where the Dirac Hamiltonian is

$$H_{\text{Dirac}} = \sigma_z \otimes \vec{\sigma} \cdot pc - \sigma_x \otimes 1mc^2. \tag{72}$$

4. QUANTUM ALGORITHMS

4.1. One Particle Quantum Simulations

4.1.1. Finite-difference Equation in 1+1 Dimensions

According to (26), a single time step of the evolution of the relativistic spinor field $\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is accomplished as follows:

$$\psi(t + \Delta t) = \mathscr{C}\mathscr{S}\psi(t). \tag{73}$$

The first step is to analyze the streaming part of (73). To do this, we write (28) as

$$\mathscr{S}\psi(z, t) = e^{-i\sigma_z \frac{p_n \Delta z}{\hbar}} \psi(z, t), \tag{74}$$

which in matrix form is

$$\mathscr{S}\begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix} \approx \begin{pmatrix} 1 - \Delta z \partial_z & 0 \\ 0 & 1 + \Delta z \partial_z \end{pmatrix} \begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix}, \tag{75}$$

where the momentum operator is taken to be $p_n \to -i\hbar\partial_z$. Equation (75) as written is exact only in the continuum limit. It follows that a consistent definition of the streaming operator for the discrete lattice is:

$$\mathscr{S}\begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix} \equiv \begin{pmatrix} \alpha(z - \Delta z, t) \\ \beta(z + \Delta z, t) \end{pmatrix}. \tag{76}$$

Taylor expanding the R.H.S. of (76) about z and keeping only the first-order terms, we recover (75). There is a geometric interpretation of

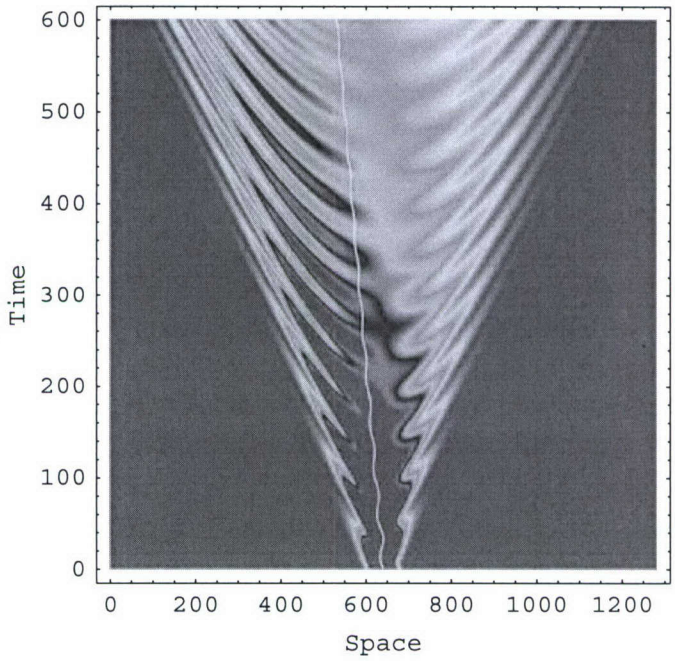


Fig. 6. Simulation of one relativistic quantum particle with $m=1, \hbar=1, \Delta z=1$ and $\Delta t=1$. Plotted is the sum of the moduli squared of the spinor components, $|\alpha|^2+|\beta|^2$, using (77). The average location of the quantum particle oscillates in time (white curve).

(76) where the α component of the spinor field at node z streams (or hops) to the neighboring node on the "right" at node $z-\Delta z$ and the β component streams to node $z+\Delta z$ on the "left," as depicted in Fig. 3b.

The second step is to analysis the collision part of (73). Substituting (27) and (76) into (73), we have

$$\begin{pmatrix}\alpha(z,t+\Delta t)\\ \beta(z,t+\Delta t)\end{pmatrix}=\begin{pmatrix}\cos\left(\frac{mc^2\Delta t}{\hbar}\right) & i\sin\left(\frac{mc^2\Delta t}{\hbar}\right)\\ i\sin\left(\frac{mc^2\Delta t}{\hbar}\right) & \cos\left(\frac{mc^2\Delta t}{\hbar}\right)\end{pmatrix}\begin{pmatrix}\alpha(z-\Delta z,t)\\ \beta(z+\Delta z,t)\end{pmatrix}. \quad (77)$$

An example simulation using (77) is shown in Fig. 6. Simulations of this sort have been carried out by Thaller [8]. Taylor expanding the collision operator to first-order in Δt, we have

$$\begin{pmatrix}\alpha(z,t+\Delta t)\\ \beta(z,t+\Delta t)\end{pmatrix}=\begin{pmatrix}1 & i\Delta t\,\frac{mc^2}{\hbar}\\ i\Delta t\,\frac{mc^2}{\hbar} & 1\end{pmatrix}\begin{pmatrix}\alpha(z-\Delta z,t)\\ \beta(z+\Delta z,t)\end{pmatrix}. \quad (78)$$

Upon multiplying out the R.H.S. and Taylor expanding about z and t, to first-order, we have

$$\Delta t\,\partial_t \begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix} = -\sigma_z \Delta z\,\partial_z \begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix} + i\,\Delta t \frac{mc^2}{\hbar} \sigma_x \begin{pmatrix} \alpha(z,t) \\ \beta(z,t) \end{pmatrix}, \tag{79}$$

which is the Dirac equation in one spatial dimension

$$i\hbar\partial_t \psi(z,t) = \left(\sigma_z pc - mc^2\sigma_x\right) \psi(z,t). \tag{80}$$

We can write (80) in propagator form

$$\psi(t) = e^{-i\hat{H}t/\hbar}\psi(0), \tag{81}$$

where the Hamiltonian is

$$\hat{H} = \sigma_z pc - mc^2\sigma_x. \tag{82}$$

This derivation of the Dirac Hamiltonian by the quantum lattice-gas method is consistent with (32) where the Feynman path integral expressed as the following Fourier transformation:

$$K(z,t) = \frac{dz}{2\hbar} \int_{-\infty}^{\infty} dp\, e^{i\frac{pz}{\hbar}}\, e^{-i\hat{H}t/\hbar}. \tag{83}$$

4.1.2. Finite-difference Equation in Higher Dimensions

The wave function is specified on a spacetime lattice with grid spacing

$$Dc^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = 0. \tag{84}$$

The local evolution equation of motion is

$$\psi' = \mathscr{S}\mathscr{C}\psi \tag{85a}$$

$$= \mathscr{S}_x\mathscr{S}_y\mathscr{S}_z\mathscr{C}\psi \tag{85b}$$

$$= e^{\sum_{i=1}^{3} \sigma_z\otimes\sigma_i \Delta r\partial_i} e^{i\frac{mc^2}{\hbar}\Delta t\sigma_x\otimes\mathbf{1}}\psi, \tag{85c}$$

which approximates the Dirac equation in the continuum limit and in the relativistic limit where $\hbar\omega \sim mc^2$ and $\hbar k \sim mc$.

The 4×4 matrix

$$\sigma_z \otimes \sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{86}$$

operating with the z spatial derivative in (85c) is diagonal whereas the matrices $\sigma_z \otimes \sigma_x$ and $\sigma_z \otimes \sigma_y$ for the x and y partial derivatives, respectively, are not diagonal.

We would like to transform $\mathscr{S}$ in (85) in such a way that all the matrices operating with the spatial partial derivatives are diagonal (and hence correspond to shifting qubits along the orthogonal lattice directions). To do this, we need the two identities:

$$e^{-i\frac{\pi}{4}\sigma_x} e^{\varepsilon\sigma_z} e^{i\frac{\pi}{4}\sigma_x} = e^{\varepsilon\sigma_y} \qquad e^{i\frac{\pi}{4}\sigma_y} e^{\varepsilon\sigma_z} e^{-i\frac{\pi}{4}\sigma_y} = e^{\varepsilon\sigma_x}, \tag{87}$$

that follow from $e^{i\frac{\pi}{4}\sigma_i} = \frac{1}{\sqrt{2}}(1 + i\sigma_i)$. Then, using the identity $\mathbf{1} \otimes e^{i\theta a} = e^{i\theta \mathbf{1}\otimes a}$, the 2-spinor similarity transformations (87) can be generalized to 4-spinor transformatons

$$\left(\mathbf{1} \otimes e^{-i\frac{\pi}{4}\sigma_x}\right) e^{\varepsilon\sigma_z\otimes\sigma_z} \left(\mathbf{1} \otimes e^{i\frac{\pi}{4}\sigma_x}\right) = e^{\varepsilon\sigma_z\otimes\sigma_y} \tag{88a}$$

$$\left(\mathbf{1} \otimes e^{i\frac{\pi}{4}\sigma_y}\right) e^{\varepsilon\sigma_z\otimes\sigma_z} \left(\mathbf{1} \otimes e^{-i\frac{\pi}{4}\sigma_y}\right) = e^{\varepsilon\sigma_z\otimes\sigma_x}, \tag{88b}$$

which we will use to diagonalize the x and y components of $\mathscr{S}$. Using (88) and defining

$$\mathscr{X}_\theta^{(1)} \equiv e^{i\theta\sigma_x} \otimes \mathbf{1} = \begin{pmatrix} \cos\theta & 0 & i\sin\theta & 0 \\ 0 & \cos\theta & 0 & i\sin\theta \\ i\sin\theta & 0 & \cos\theta & 0 \\ 0 & i\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{89a}$$

$$\mathscr{X}_\theta^{(2)} \equiv \mathbf{1} \otimes e^{i\theta\sigma_x} = \begin{pmatrix} \cos\theta & i\sin\theta & 0 & 0 \\ i\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & i\sin\theta \\ 0 & 0 & i\sin\theta & \cos\theta \end{pmatrix}, \tag{89b}$$

and

$$\mathscr{Y}_\theta^{(2)} \equiv \mathbf{1} \otimes e^{i\theta\sigma_y} = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix}, \tag{89c}$$

and

$$S = S_x S_y S_z \equiv e^{\sigma_z\otimes\sigma_z \Delta r \nabla}, \tag{90}$$

where $S_w \equiv e^{\sigma_z \otimes \sigma_z \Delta r \partial_w}$, the product components of $\mathscr{S} = \mathscr{S}_x \mathscr{S}_y \mathscr{S}_z$ can be written

$$\mathscr{S}_x = e^{\sigma_z \otimes \sigma_x \Delta r \partial_x} = \mathscr{Y}^{(2)}_{\frac{\pi}{4}} S_x \mathscr{Y}^{(2)\dagger}_{\frac{\pi}{4}} \tag{91a}$$

$$\mathscr{S}_y = e^{\sigma_z \otimes \sigma_y \Delta r \partial_y} = \mathscr{X}^{(2)\dagger}_{\frac{\pi}{4}} S_y \mathscr{X}^{(2)}_{\frac{\pi}{4}} \tag{91b}$$

$$\mathscr{S}_z = e^{\sigma_z \otimes \sigma_z \Delta r \partial_x} = S_z, \tag{91c}$$

so (85) can be rewritten as the following quantum algorithm:

$$\psi' = \mathscr{Y}^{(2)}_{\frac{\pi}{4}} S_x \mathscr{Y}^{(2)\dagger}_{\frac{\pi}{4}} \mathscr{X}^{(2)\dagger}_{\frac{\pi}{4}} S_y \mathscr{X}^{(2)}_{\frac{\pi}{4}} S_z \mathscr{X}^{(1)}_{\frac{mc^2 \Delta t}{\hbar}} \psi. \tag{92}$$

The quantum algorithm (92) has local qubit-qubit interaction operators $\mathscr{X}^{(2)}_{\frac{\pi}{4}}$, $\mathscr{Y}^{(2)}_{\frac{\pi}{4}}$ and $\mathscr{C} = \mathscr{X}^{(1)}_{\frac{mc^2 \Delta t}{\hbar}}$ (quantum collision operators), as well as cubic lattice translation operators S_x, S_y, and S_z (classical stream operators). It is demonstrated in 2+1 dimensions for a single massive Dirac particle, see Fig. 7.

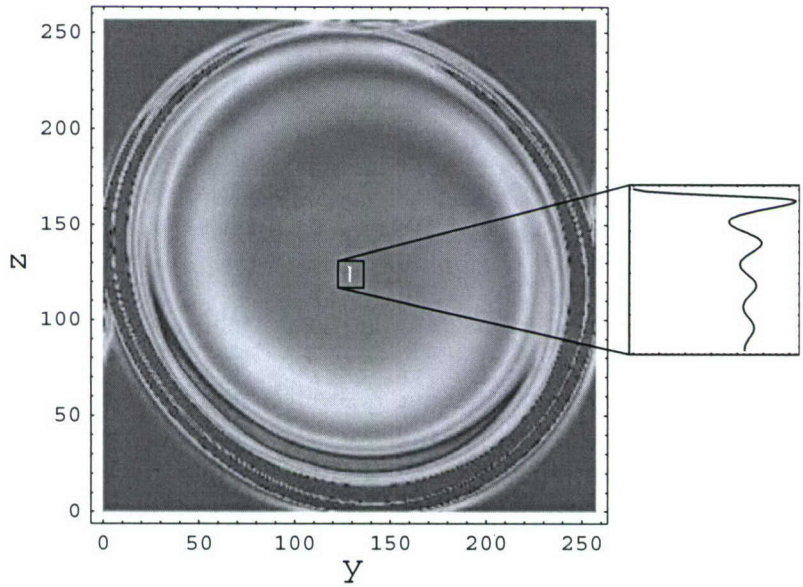


Fig. 7. Simulation of one relativistic quantum particle with $m = 1, \hbar = 1$ in 2+1 dimensions ($\Delta y = 1$, $\Delta z = 1$ and $\Delta t = 1$). Plotted is the sum of the moduli squared of the spinor components, $|\alpha|^2 + |\beta|^2$, at time $t = 128\Delta t$. The mean value of the position of the quantum particle oscillates in time (white curve in center and black curve in the expanded view on the right).

For numerical purposes in the one particle case, the discrete amplitude field is:

$$\psi(\vec{x},t)=\begin{pmatrix}\alpha(\vec{x},t)\\ \beta(\vec{x},t)\\ \mu(\vec{x},t)\\ \nu(\vec{x},t)\end{pmatrix}, \tag{93}$$

where the spacetime point $(\vec{x},t)$ coincides with the nodes of a hypercubic lattice with grid size Δr. The operators S_w induce a finite displacement

$$S_w\psi(\vec{x})=\psi\left(\vec{x}+\Delta r(\sigma_z\otimes\sigma_z)_{ww}\hat{w}\right), \tag{94}$$

of the components of the 4-spinor only along lattice directions:

$$S_x\psi(x,y,z)=\begin{pmatrix}\alpha(x+\Delta r,y,z)\\ \beta(x-\Delta r,y,z)\\ \mu(x-\Delta r,y,z)\\ \nu(x+\Delta r,y,z)\end{pmatrix}, \tag{95a}$$

$$S_y\psi(x,y,z)=\begin{pmatrix}\alpha(x,y+\Delta r,z)\\ \beta(x,y-\Delta r,z)\\ \mu(x,y-\Delta r,z)\\ \nu(x,y+\Delta r,z)\end{pmatrix}, \tag{95b}$$

and

$$S_z\psi(x,y,z)=\begin{pmatrix}\alpha(x,y,z+\Delta r)\\ \beta(x,y,z-\Delta r)\\ \mu(x,y,z-\Delta r)\\ \nu(x,y,z+\Delta r)\end{pmatrix}. \tag{95c}$$

These streaming operators are data shifting operators causing permutations of the components of the 4-spinor wave function across the lattice. The collision operators act independently on each node of the lattice and cause local quantum entanglement between component pairs of the 4-spinor. The streaming operators in turn propagate this local on-site entanglement to next nearest neighbors so that eventually quantum entanglement covers the entire lattice.

It is possible to rewrite (92) as a finite difference equation on a body-centered cubical lattice. The resulting set of coupled finite difference equations are similar to the finite difference representation of the 3D Dirac equation given by Bialynicki-Birula in 1994.[9] A drawback of expressing the algorithm as a finite-difference equation is its unsuitability

for a quantum computer implementation using two-qubit quantum gates whereas our manifestly unitary expression (92) is suitable.

A continuous effective field theory for $\psi = (\alpha, \beta, \mu, \nu)$ follows in the continuum limit of the emergent finite-difference equations by Taylor expanding in the small grid sizes. We obtain

$$\partial_t \begin{pmatrix} \alpha \\ \beta \\ \mu \\ \nu \end{pmatrix} = \frac{\Delta r}{\Delta t}\partial_x \begin{pmatrix} -\beta \\ -\alpha \\ \nu \\ \mu \end{pmatrix} + i\frac{\Delta r}{\Delta t}\partial_y \begin{pmatrix} \beta \\ -\alpha \\ -\nu \\ \mu \end{pmatrix} + \frac{\Delta r}{\Delta t}\partial_z \begin{pmatrix} \alpha \\ -\beta \\ -\mu \\ \nu \end{pmatrix} - i\frac{mc^2}{\hbar} \begin{pmatrix} \mu \\ \nu \\ \alpha \\ \beta \end{pmatrix} + \mathcal{O}(c\Delta r, \Delta t), \tag{96}$$

which is exactly the Dirac equation

$$\partial_t \psi = c \sum_i \sigma_z \otimes \sigma_i \partial_i \psi - i\sigma_x \otimes 1 \frac{mc^2}{\hbar}\psi, \tag{97}$$

when $\Delta t \sim \Delta r \sim \varepsilon$ are infinitesimal and when the partial derivative with respect to time is defined as $\partial_t \psi \equiv \lim_{\Delta t \to 0} \frac{\psi' - \psi}{\Delta t}$. Equation (92) gives rise to perfectly unitary evolution of the discretized wave function and, therefore, is an unconditionally stable numerical algorithm. The quantum algorithm (92) for modeling (96) is less than 1st-order convergent; see Fig. 8. For practical numerical purposes, we will need to modify (92) to improve the algorithm's convergence properties.

4.1.3. Improved Finite-difference Equation in 3+1 Dimensions

Our basic approach to improving the accuracy of the quantum algorithm is to set the grid size Δr to be smaller than the Compton wavelength $\lambda = \frac{h}{mc}$ of the modeled particle

$$\Delta r \sim \varepsilon \frac{h}{mc}, \tag{98}$$

and to introduce a small temporal scale that is much smaller than $\frac{\lambda}{c}$

$$\Delta t \sim \varepsilon^2 \frac{h}{mc^2}. \tag{99}$$

The diffusive ordering condition of spatial and temporal fluctuations typical of random walk processes, $\Delta r^2 = \nu \Delta t$, provides a context to understand the scaling behavior of the small parameter ε. According to (98) and

(99), the diffusive transport coefficient is $\nu = \frac{h}{\varepsilon m}$ and the unit lattice velocity is $\frac{\Delta r}{\Delta t} = \frac{c}{\varepsilon}$, which approaches infinity as $\varepsilon \to 0$. In this limit, the mean velocity of the modeled quantum particle is relatively small, compared to $\Delta r/\Delta t$, hence the resulting effective field theory corresponds to the nonrelativistic limit of the Dirac equation as $\varepsilon \to 0$.

To diagonalize the stream operators in (91), we used a fixed and finite rotation angle $\frac{\pi}{4}$ independent of the grid resolution. We will now diagonalize the stream operators using a small rotation angle proportional to Δt. By (99), the rotation angle is $\theta = \frac{mc^2 \Delta t}{h} = \varepsilon^2$, which is dependent on the grid resolution. The displacement operators in the Dirac equation can be represented by interleaving stream and collision operators on a cubical lattice as follows:

$$\mathscr{S}_x = e^{\sigma_z \otimes \sigma_x \Delta r \partial_x} \to S_{-x}^{2,4} \mathscr{Y}_{\frac{\varepsilon}{2}}^{(2)} S_x^{2,4} \mathscr{Y}_{\frac{\varepsilon}{2}}^{(2)\dagger} S_x^{1,3} \mathscr{Y}_{\frac{\varepsilon}{2}}^{(2)} S_{-x}^{1,3} \mathscr{Y}_{\frac{\varepsilon}{2}}^{(2)\dagger} \tag{100}$$

and

$$\mathscr{S}_y = e^{\sigma_z \otimes \sigma_y \Delta r \partial_y} \to S_{-y}^{2,4} \mathscr{X}_{\frac{\varepsilon}{2}}^{(2)\dagger} S_y^{2,4} \mathscr{X}_{\frac{\varepsilon}{2}}^{(2)} S_y^{1,3} \mathscr{X}_{\frac{\varepsilon}{2}}^{(2)\dagger} S_{-y}^{1,3} \mathscr{X}_{\frac{\varepsilon}{2}}^{(2)}, \tag{101}$$

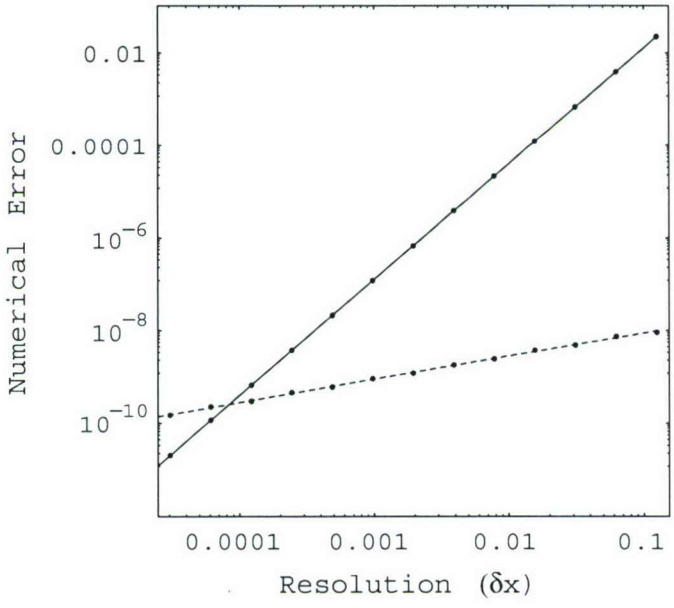


Fig. 8. L2 norm error $\sqrt{\frac{1}{L}\sum_{i=1}^{L}[|\psi(x_i)|^2 - |\psi_{\text{ex}}(x_i)|^2]}$ plotted versus grid resolution $\Delta x = \frac{1}{L}$ for numerical simulations of the 1D Weyl equation (Dirac equation with $m = 0$) with lattice sizes from $L = 8$ to $L = 32768$. The error curve's slope of the original and improved algorithm is 0.5 (dashed line) and 2.5 (solid line), respectively. This demonstrates the high numerical accuracy of the improved (symmetrized and interleaved) quantum algorithm.

where the superscripts on the streaming operators refer to individual components of the 4-spinor. For example, the classical streaming operators in (95) are $S_w = S_{-w}^{2,3} S_w^{1,4}$. The quantum stream operators (91a) and (91b) are now redefined in terms of the respective component-wise classical stream operators separated by collision operators according to (100) and (101). This kind of interleaving of streaming and collision operators removes the spurious check-board effect of independent sublattice dynamics that otherwise occurs.[10,11] So far we treated the non-diagonal quantum stream operators $e^{\sigma_z \otimes \sigma_x \Delta r \partial_x}$ and $e^{\sigma_z \otimes \sigma_y \Delta r \partial_y}$, but not the classical operator $e^{\sigma_z \otimes \sigma_z \Delta r \partial_z}$ because no such improvement exists since it is diagonal. However, if instead of using the Dirac matrix $\sigma_z \otimes \sigma_z$, we use an alternative non-diagonal representation for the z-direction partial derivative, then we can again employ interleaving. Therefore, we consider the following alternate form of the Dirac equation:

$$\partial_t \psi = c\sigma_z \otimes \sigma_x \partial_x \psi + c\sigma_z \otimes \sigma_y \partial_y \psi + c\sigma_y \otimes \mathbf{1} \partial_z \psi - i\sigma_x \otimes 1 \frac{mc^2}{\hbar} \psi. \quad (102)$$

Now the quantum stream operator in (102) for the z-direction can be re-expressed in a fashion similar to (100) and (101) as

$$\mathscr{S}_z \equiv e^{\sigma_y \otimes \mathbf{1} \Delta r \partial_z} \to S_z^{2,3} \mathscr{X}_{\frac{\varepsilon}{2}}^{(1)} S_{-z}^{2,3} \mathscr{X}_{\frac{\varepsilon}{2}}^{(1)\dagger} S_z^{1,4} \mathscr{X}_{\frac{\varepsilon}{2}}^{(1)} S_{-z}^{1,4} \mathscr{X}_{\frac{\varepsilon}{2}}^{(1)\dagger}. \quad (103)$$

Then instead of (91), we use (100), (101), and (103) for an improved quantum algorithm

$$\psi(t + \Delta t) = \mathscr{S}_x \mathscr{S}_y \mathscr{S}_z \psi(t). \quad (104)$$

In the interleaved algorithm (104), the mass term in the Dirac equation is automatically produced at order ε^2.

It is possible to derive a finite-difference equation representation of the quantum lattice-gas algorithm (104) by carrying out all the collision and streaming operations symbolically. The result expressed as a finite-difference equation is no longer expressible using a stencil including only the nearest neighbors of the body-centered cubic lattice. Nevertheless, once again, a continuous effective field theory, in the one-particle sector of the Hilbert space, for $\psi = (\alpha, \beta, \mu, \nu)$ follows in the continuum limit as we

Taylor expand in Δr and in Δt:

$$\partial_t \begin{pmatrix} \alpha \\ \beta \\ \mu \\ \nu \end{pmatrix} = \frac{\Delta r}{\Delta t}\partial_x \begin{pmatrix} -\beta \\ -\alpha \\ \nu \\ \mu \end{pmatrix} + i\frac{\Delta r}{\Delta t}\partial_y \begin{pmatrix} \beta \\ -\alpha \\ -\nu \\ \mu \end{pmatrix} + i\frac{\Delta r}{\Delta t}\partial_z \begin{pmatrix} -\mu \\ -\nu \\ \alpha \\ \beta \end{pmatrix} + i\frac{mc^2}{\hbar} \begin{pmatrix} \mu \\ \nu \\ \alpha \\ \beta \end{pmatrix} + \mathcal{O}(\Delta r^4, \Delta t^2), \tag{105}$$

which approximates the Dirac equation (102) when Δt is small. In (105), according to symbolic mathematics, the mass term arises at order ε^2, the spatial terms arise at order ε^3, and the error terms arise at order ε^4. Yet, according to numerical simulations, the effective field theory (105) is 1st-order convergent. With the quantum stream operator $\mathcal{S} = \mathcal{S}_x \mathcal{S}_y \mathcal{S}_z$, we define the dual operator $\tilde{\mathcal{S}} \equiv \mathcal{S}^\dagger_{-x} \mathcal{S}^\dagger_{-y} \mathcal{S}^\dagger_{-z}$, by taking the adjoint of the quantum stream operators and reversing the displacement directions. Then, it is possible use a symmetrized evolution operator[11]

$$\psi(t + \Delta t) = \tilde{\mathcal{S}} \mathcal{S} e^{-\zeta \Delta t^2} \psi(t), \tag{106}$$

where ζ is a constant parameter. According to numerical simulation, the interleaved and symmetrized algorithm (106) is better than 2nd-order convergent, as demonstrated in Fig. 8 for the 1+1 dimensional case.

4.2. Many Particle Quantum Simulations

A useful characteristic of the quantum algorithm to model a many-body system of Dirac particles is that it uses the same protocol for local collisions and local translations used to model the dynamical behavior of the one-body problem governed by the Dirac equation. Therefore, the computational overhead associated with the action of the unitary operators is fixed for both the one-body and many-body cases. What differs is the number of quantum states (configurations of qubits) that must be dealt with in the numerical simulation.

To model all the dynamics of many Dirac particles, we use an interfering set of probability amplitudes defined on the lattice. All these probability amplitudes, considered as an ordered finite set of complex numbers, constitutes a discrete amplitude field with 2^Q complex components for a quantum computer with Q qubits. With 4 qubits per node on a body-centered cubic lattice with $V = L^3$ nodes, the quantum state of the

system is the following tensor product:

$$|\Psi(t)\rangle = \bigotimes_{q=1}^{4} \bigotimes_{x=1}^{L} \bigotimes_{y=1}^{L} \bigotimes_{z=1}^{L} |q(x, y, z, t)\rangle. \tag{107}$$

The quantum algorithm amounts to unfolding a quantum mechanical evolution operator into an ordered sequence on 2-qubit quantum gate operations. We have analytically determined an accurate approximation of the local Hamiltonian, say $\mathscr{H}_n$ for the nth lattice site, finding it to be the Dirac Hamiltonian $\sigma_z \otimes \vec{\sigma} \cdot \vec{p}c + \sigma_x \otimes \mathbf{1}mc^2$ with momentum $\vec{p} = (p_x, p_y, p_z)$ and the Pauli spin matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The Dirac Hamiltonian is an effective local Hamiltonian. The total system Hamiltonian $\mathcal{H}$ formally generates the quantum mechanical evolution:

$$\mathscr{U} = e^{-i\mathcal{H}\Delta t/\hbar} = \bigotimes_{n=0}^{N} e^{-i\mathscr{H}_n \Delta t/\hbar}, \tag{108}$$

represented by a $2^Q \times 2^Q$ matrix. An analytical expression of $\mathcal{H}$ could have an infinite number of terms due to the Campbell-Baker-Hausdorff theorem and the non-commutability of the quantum gates ($[\mathscr{H}_n, \mathscr{H}_m] \neq 0$). Nevertheless, this is a crucial feature of the quantum algorithm: it is possible to numerically compute an accurate approximation of $\mathcal{H}$, a *many-body Dirac Hamiltonian*, as a total system generator of the spacetime evolution. It is in this sense that the quantum algorithm is isomorphic to a Feynman path summation.

In general, the quantum mechanical evolution is expressed as a Schrödinger propagation equation as follows:

$$|\Psi(t+\tau)\rangle = \mathcal{U}|\Psi(t)\rangle. \tag{109}$$

Equation (109) has an implied operator ordering that must be retained in the quantum algorithmic representation. $\mathcal{U}$ depends on the local unitary operator $\mathscr{U} = \mathscr{C}\mathscr{S}$, which changes the probability amplitudes on each node independently. So the L^3-fold tensor product over the lattice nodes is the total quantum system evolution operator:

$$\mathcal{U} = \bigotimes_{\ell=1}^{L^3} \mathscr{C}\mathscr{S}. \tag{110}$$

Therefore, the specifying the many-particle form of the quantum algorithms (92) and (106) entails representing $\mathscr{S}$ and $\mathscr{C}$ in (110) in terms of

conservative and local 2-qubit quantum gates. This is accomplished using fermionic ladder operators in a second quantized representation.

With $\hat{a}_\alpha^\dagger, \hat{a}_\alpha$, and $\hat{n} = \hat{a}_\alpha^\dagger \hat{a}_\alpha$ denoting the fermionic creation, annihilation, and number operator, respectively, of the αth qubit ($1 \le \alpha \le 4L^3$), the collision operators (89) are the following:

$$\mathscr{X}_{\alpha\beta}(\theta) = \mathbf{1} - i \sin\theta(\hat{a}_\alpha^\dagger \hat{a}_\beta + \hat{a}_\beta^\dagger \hat{a}_\alpha) + (\cos\theta - 1)(\hat{n}_\alpha + \hat{n}_\beta) - 2\cos\theta \hat{n}_\alpha \hat{n}_\beta \tag{111a}$$

$$\mathscr{Y}_{\alpha\beta}(\theta) = \mathbf{1} + \sin\theta(\hat{a}_\alpha^\dagger \hat{a}_\beta - \hat{a}_\beta^\dagger \hat{a}_\alpha) + (\cos\theta - 1)(\hat{n}_\alpha + \hat{n}_\beta) - 2\cos\theta \hat{n}_\alpha \hat{n}_\beta, \tag{111b}$$

where α and β index different qubits at the same site.[11] Then (89) are rewritten as

$$\mathscr{X}_\theta^{(1)} \to \mathscr{X}_{13}(\theta)\mathscr{X}_{24}(\theta) \tag{112a}$$

$$\mathscr{X}_\theta^{(2)} \to \mathscr{X}_{12}(\theta)\mathscr{X}_{34}(\theta) \tag{112b}$$

$$\mathscr{Y}_\theta^{(2)} \to \mathscr{Y}_{12}(\theta)\mathscr{Y}_{34}(\theta). \tag{112c}$$

$2L^3$ applications of either $\mathscr{X}_{\alpha\beta}$ or $\mathscr{Y}_{\alpha\beta}$ are required for a single collision step.

Streaming occurs by successive application of the interchange operator[12]

$$\hat{S}_{\mu\nu} = \mathbf{1} + \hat{a}_\mu^\dagger \hat{a}_\nu + \hat{a}_\nu^\dagger \hat{a}_\mu - \hat{n}_\mu + \hat{n}_\nu. \tag{113}$$

$(L-1)^3$ number of applications of $\hat{S}_{\mu\nu}$ (μ refers to one qubit-component at some site and ν to the same component at its neighboring site) are required to stream each qubit along a cubic lattice direction.

The total evolution operator $\hat{E}$ is the product of collision operators $\mathscr{X}$ and $\mathscr{Y}$ and streaming operator S corresponding to algorithm (92) or some variant of (104) depending on the desired degree of numerical accuracy.

With the new ket $|\Psi'(t+\Delta t)\rangle = \hat{E}|\Psi(t)\rangle$, the resulting probability of finding a particle at site $\vec{x}$ is

$$P(\vec{x}) = \sum_{i=1}^{4} \langle \Psi' | \hat{n}_{\alpha+i} | \Psi' \rangle, \tag{114}$$

where α the index of the 1st qubit at $\vec{x}$.

With 4 qubits per node, the local propagator $\mathscr{U}$ is represented by a 16×16 size matrix. This appears to be larger than necessary, because there are only 4 component to a Dirac spinor. However, for the purpose of

being able to model a many-body quantum system, we require one qubit for each component of the Dirac spinor.

One may strive to execute the maximum possible number of quantum gates in parallel to save physical time in the actual implementation because of the experimental limitations due to de-phasing related to spin-spin decoherence and spin relaxation. The issue of operator ordering is handled intrinsically by the quantum algorithm in the case when quantum gates are applied in parallel to different and non-overlapping pairs of qubits (see Table 1).

4.2.1. Algorithmic Complexity

To quantify the algorithmic complicity associated with evolving the modeled quantum wave function by one unit time Δt, we count the total number of required basic quantum logic operations. We count each conservative collide operator, $\mathscr{X}$ and $\mathscr{Y}$ defined in (111), as one basic quantum logic operation. So according to (112), the quantum mechanical collide operators each take 2 basic quantum logic operations. Furthermore, we count the operation (113) of streaming a single qubit between neighboring nodes as one basic quantum logic operation. That is, streaming a single component of the Dirac spinor along one orthogonal lattice direction (cubic lattice vector) counts as one quantum logic operation.

The algorithmic complexity for numerical simulations a system of Dirac particles in 3+1 dimensions for one time step scales as $C = 2\rho_c L^3 + \rho_s (L-1)^3$, where ρ_c and ρ_s are the number of operations per node for collisions and streaming, respectively. This measure of the algorithmic complexity counts the total number of quantum gate operations required to updated the modeled wave function, whether or not those operations are performed in a serial or parallel fashion. For the simplest algorithm (92), $\rho_c = 5$ and $\rho_s = 12$, and for the improved algorithm (104), $\rho_c = 12$ and $\rho_s = 24$. Both ρ_c and ρ_s double when we use a symmetrized rule like (106) but are a fixed-cost overhead. With $Q = 4L^3$ qubits, the size of

Table 1. Summary of the number of local quantum gate operation per node

Algorithmic variant	Eq. No.	$2\rho_c$	ρ_s	Total Ops.
Standard	(92)	10	12	22
Interleaved	(104)	24	24	48
Symmetrized interleaved	(106)	49	48	97

the Hilbert space is exponential 2^Q, whereas the complexity $C = \frac{\rho_c}{2} Q + \rho_s [Q - \frac{3}{4}(2Q)^{\frac{2}{3}} + \frac{3}{2}(2Q)^{\frac{1}{3}} - 1]$ for all the variants of our quantum algorithm is dominantly linear in Q.

5. CONCLUSION

5.1. Summary

We have presented a solution of Feynman's chessboard problem in 3+1 dimensions. With this result, we have shown how to design a quantum algorithm suitable for implementation on a quantum computer to accurately model one or more Dirac particles. In the one-body case, the mechanical wave equation is the Dirac equation (or Weyl equation in the case when $m = 0$). We say the quantum computer performs analog computation because it relies on one physical quantum mechanical system (a system of qubits with only local qubit-qubit interaction) to model another physical quantum mechanical system, a relativistic Dirac system. Each qubit contained within the quantum computer is embodied by a two-level quantum system, such as a spin-$\frac{1}{2}$ nucleus. Every quantum logic gate operation between local qubit pairs may be embodied by a local interaction Hamiltonian, say a secular dipolar Hamiltonian for the spin–spin coupling.

The initial state of a quantum node is an unequally weighted superposition of states over the Hilbert space of that node, generated by tipping, or rotating, each qubit of the quantum node independently. The quantum algorithm approximates a spatially continuous wave function using a numerical grid with finite resolution, and that provides a natural cut-off. From the point-of-view of the modeler, there exists a small numerical grid-level scale below the physical microscopic scale of the modeled quantum mechanical system.

Through the combined actions of the unitary collision and translation operators, all the quantum mechanical pathways of the possible motions of the modeled quantum particles are numerically treated. The numerical model can be viewed as a kind of kinetic system of locally interacting spins on the small grid-level scale. There also exists a large scale, which corresponds to the long wavelength limit of the dynamical modes in the discrete spin system. This numerical large scale is equivalent to the physical microscopic scale where the many-body wave function of the modeled quantum mechanical system in question is well defined. A continuous wave function is accurately approximated as one approaches the continuum limit where the grid resolution of the spatial lattice become infinite (the lattice cell size

approaches zero). In the long wave length limit, the dynamical amplitude field is both continuous and differentiable in space and time. An emergent effective field theory, such as (96) or (105), is analytically determined by Taylor expanding a finite-difference equation representation of the dynamics. What is physical in the model is the finite-difference equation. That is, it is this finite-difference equation that governs the dynamical amplitude field. The effective field theory, such as (96) or (105), is actually a partial differential equation of infinite order. Only the low-order time and space partial derivatives are numerically relevant during the early stage of evolution of the wave function and all the higher order partial derivatives are considered error terms during this early stage.

5.2. Final Remarks

The method can be used to analytically derive the relativistic Hamiltonian for a small number of Dirac particles. For example, the few-body Hamiltonians, and the associated coupled set of partial differential equations of motion, will be presented in a subsequent paper. In any case, to be a practical and useful algorithm, it must have at least second-order convergence which means that decreasing the grid cell size by a factor of two causes the numerical error inherent in the algorithm to decrease by a factor of four. We have demonstrated it is possible to improve the numerical convergence of the quantum algorithmic method to meet this requirement.

All the qubits in the quantum computer must remain phase-coherent, i.e. globally entangled, over the entire course of the quantum computation. In any practical experimental implementation, this quantum algorithm may therefore require the use of extra qubits to correct for bit-flip and phase errors,[13,14] significantly increasing the number of required qubits per node.

Quantum logical operations are embodied in the spin–spin coupling mechanism, or some other mathematically equivalent two-qubit coupling mechanism appropriate to the particular experimental realization. The form of a naturally occurring interaction Hamiltonian can be effectively altered by applying appropriately timed spin rotations using externally applied electromagnetic pulses, whose frequency, strength, polarization, and pulse duration depends on the particular experimental realization, so that the natural coupling along with the program of externally applied pulses together cause the desired quantum lattice-gas collision operation. [15–17]5

---

[5] A subsequent paper will address how to use nuclear magnetic resonance spectroscopy of a solid-state crystal with a secular dipolar Hamiltonian for two spin-$\frac{1}{2}$ nuclei to emulate both the quantum lattice-gas collision and stream operators for numerically predicting space-time dependent solutions for the wave function of a system of Dirac particles.

In relativistic formuatlions one typically sees space and time as intrinsically coupled, yet the quantum algorithm presented here based on local stream and collide operators rigidly seems to enforce a Galilean-like separation between space and time-space and time are geometrically coupled in a regular lattice structure. Yet since the fundamental stream and collide operators do not commute, the continuous Lorentz transformations of special relativity emerge in the long-wave length limit. Furthermore, the dynamical equation of motion governing the continuous four spinor field modeled by this method remains invariant under the Lorentz transformation, but only in the continuum limit. The four spinor field may be treated as a continuous Dirac field in the case when the lattice has infinite resolution (this is, when the lattice cell size is infinitesimal). So any Galilean-like separation can occur only at a very small scale. If the operative relativistic quantum mechanical evolution is generated by the Dirac Hamiltonian, leading to an effective field theory perfectly first order in its space and time derivatives, then the small scale (lattice cell size) must be infinitesimal, and hence it would always be beyond experimental detectably. However, if a discrete sub-lattice were a physical property of spacetime, and it had a finite cell size (perhaps on the order of the Planck scale), then there may be physical significance to the hyper bcc spacetime lattice. It remains an outstanding open problem to determine whether or not it is possible experimentally to measure any lattice property of spacetime, such as a fundamental grid size.

ACKNOWLEDGMENTS

I would like to thank Timothy Havel for pointing me to useful references and for proofreading this manuscript. I would especially like to express my indebtedness to the Air Force Office of Scientific Research for their support of quantum computational physics research.

REFERENCES

1. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, 1965). Problem 2–6 on page 34.
2. R. P. Feynman, Geometry of dirac equ. in 1 dimension. *California Institute of Technology CIT archives*, (13.3), February 1946.
3. T. Jacobson and L. S. Schulman, *J. Phys. A: Math. Gen.* **17**, 375 (1984).
4. T. Jacobson, *J. Phys. A: Math. Gen.* **17**, 2433 (1984).
5. B. M. Boghosian and W. Taylor IV, *Phys. Rev. E* **57**, 54 (1998).
6. R. P. Feynman, *Rev. Modern Phys.* **20**(2), 367 (1948).

7. S. S. Schweber, *Rev. Modern Phys.* **449**(2), 449 (1986).
8. B. Thaller, Visualizing the kinematics of relativistic wave packets. *arXiv:quant-ph*, (0409079):7, 2004.
9. I. Bialynicki-Birula, *Phys. Rev. D* **49**(12), 6920 (1994).
10. J. Yepez, *Int. J. Modern Phys. C* **12**(9), 1285 (2001).
11. J. Yepez and B. Boghosian, *Comput. Phys. Commun.* **146**(3), 280 (2002).
12. J. Yepez, *Phys. Rev. E* (046702), 18 (2001).
13. P. W. Shor, *Phys. Rev. A* **52**, R2493 (1995).
14. E. Knill, R. Laflamme, and W. H. Zurek, *Science*, **279**, 342 (1998).
15. M. Pravia, Z. Chen, J. Yepez, and D. G. Cory, *Comput. Phys. Commun.* **146**(3), 339 (2002).
16. M. Pravia, Z. Chen, J. Yepez, and D. G. Cory, *Quant. Informat. Process.* **2**, 1 (2003).
17. Z. Chen, J. Yepez, and D. G. Cory, *Phys. Rev. A* To appear 2006. arXiv:quant-ph/0410198.